Crystallographic Studies on Structure-Function Relationships of Alkaline Protease from Pseudomonas aeruginosa IFO3080(Dissertation_全文)

AUTHOR(S):

Miyatake, Hideyuki

(2)

主論文

Crystallographic Studies on Structure-Function Relationships of Alkaline Protease from *Pseudomonas aeruginosa* IFO3080

Hideyuki Miyatake

1996

Contents

List of Tables

List of Figures

要旨

緑膿菌*Pseudomonos aeruginosa* IFO3080 由来のアルカリプロテアーゼはメトジンシンと呼ばれるファミリーに分類される亜鉛金属エンドプロテアーゼの一種である。アルカリプロテアーゼは分子量約５万、至適pHを8～9の弱塩基性側に持ち、様々なプロテアーゼ阻害剤やイミュノグロブリンの不活性化、またハーゲマンファクターの活性化などによって緑膿菌の菌体外分泌毒素として働く。このアルカリプロテアーゼの構造と機能の関係を解明するために、その立体構造をX-線結晶構造解析の手法により2.0Å分解能で決定した。本酵素は、他のメトジンシンに属する酵素の加水分解ドメインと共通のフォールディングトポロジーを持つ、N-端（加水分解）ドメインと、β-ストランド優勢なC-端ドメインから成ることが明らかになった。更に、アルカリプロテアーゼの触媒機構を解明するため１種類の合成阻害剤と２種類の基質類似ペプチドの計３種類のリガンドとの複合体をソーキング法で調整し、2.3 Åの分解能で構造解析を行った。同時に、アルカリプロテアーゼと、構造既知のメトジンシンに属する酵素との構造比較を行い、それぞれの構造と機能の違いについて検討した。これら結果、アルカリプロテアーゼのS_1'ポケットは他のメトジンシンと比べて基質のP_1'位の側鎖に対する立体障害が低いことが明らかになった。更に、各リガンドの非プライム部分はプライム部分に比べ、自由なコンフォメーションを採っていた。更に、ネイティブ構造ではフレキシブルなループ部分がリガンドの結合に伴って特定のコンフォメーションを採り、基質の主鎖と水素結合を形成しながらミカエリス複合体を活性クレフト内に閉じ込めるようにして安定化していた。これらの結果は、すでに実験的に示されていたアルカリプロテアーゼの広い基質特異性を説明する。一方、特にP_1位に塩基性の残基を持つ基質に対しては、アルカリプロテアーゼはフレキシブルループ中のGlu194の側鎖のコンフォメーションを変化させることにより、より強い親和性を示すことが示唆された。

もう一方のβストランド優勢なドメイン（C-端ドメイン）は、更に構造上の特徴によって3つのサブドメインに分けることができた。その一つは、β-ストランドが右巻ら旋状に連結されたβ-ヘリックスと呼ばれる新規な構造を採っており、その各ターンにはカルシウムイオンが配位していた。また、最もC-端に近いサブドメインには分泌のためのシグナル領域が存在することがアルカリプロテアーゼと類似の酵素との比較により示唆された。これらの考察より、C-端ドメインは、不活性なアルカリプロテアーゼ前駆体を分泌する際に必須なシグナルを、その分泌機構に対して示す役割と、カルシウムイオンが仲介する自蝕作用によって、前駆体を活性化する際に重要な役割を果たしていることが示唆された。

Abstract

Alkaline protease from *Pseudomonas aeruginosa* IFO3080 is a member of a superfamily 'metzincins'. The crystal structure of the alkaline protease was solved at 2.0 Å resolution. The enzyme consists of two domains: proteolytic (N-terminal) domain and a C-terminal domain. The proteolytic domains of the metzincins including the alkaline protease share a common folding topology. The C-terminal domain is β-strand dominant comprising eighteen β-strands. To elucidate the catalytic mechanism of the alkaline protease, complexes of the enzyme with a synthetic inhibitor and two substrate-analogue peptides, were prepared by soaking method. Furthermore, structural comparisons between the alkaline protease and other structurally known metzincins were carried out to reveal the structural differences responsible for their functions. Through these studies, it has been revealed that steric hindrance surrounding the S_1' pocket of the alkaline protease is relatively low compared with that of other metzincins. A loop region, which is flexible in the native structure of the alkaline protease, adopts a fixed conformation on ligand binding, stabilizing the Michaelis complexes making hydrogen bonds with the ligands. Each of the non-primed portions of the ligands can adopt relatively free conformations compared with those of the primed-portions. All these observations explain the broad specificity of the alkaline protease. Moreover, through a modeling study of the complexes on a graphic workstation, it is suggested that basic residues such as arginine or lysine in P_1 position of a ligand can enhance binding affinities toward the enzyme by making salt bridges with the fixed flexible loop. This observation corresponds to experimental results showing that the enzyme preferentially hydrolyzes a substrate of the C-terminal side of a basic residue. Taking into acount the structures of the native enzyme, its complexes, and the results previously obtained through the biochemical experiments, the catalytic mechanism of the enzyme is proposed.

The alkaline protease also has an additional C-terminal domain. This domain is β-strand dominant and can be further subdivided into three regions: catalytic domain proximal region, C-terminal region, and β-helix region. The C-terminal region appears to provide a signal essential for the secretion of the inactive higher molecular weight zymogen of the alkaline protease into the extracellular medium. The folding of the right-handed β-helix region may be associated with the activation process mediated by calcium ions after secretion, due to a large conformational change on the ion binding to the region. The domain may play a significant role in the maturation of the enzyme and/or secretion into the extracellular medium from the bacterium.

Chapter 1

Introduction

The elucidation of tertiary sturctures of biological substances in detail at atomic resolution has been bringing us a variety of invaluable knowledge. One brilliant example is the discovery of DNA double helical structure, which was made by J. D. Watson and F. H. C. Crick in 1953 [1, 2], by X-ray crystallography. It can be said that modern biology, termed 'molecular biology', started with this discovery because it led the way to understanding life phenomena in molecular terms. Since this success, the observation of biological substances at atomic resolution, especially using X-ray crystallography, has become one of the most powerful ways to understand living organisms. Accumulated evidence has shown us that macromolecules, such as proteins, nucleic acids, lipids and multi-macromolecular assemblies adopt their characteristic structures *in vivo*, and these conformations are strongly associated with their functions. Nowadays, to know the tertiary structures of macromolecules is a significant and essential step to discuss the functions of macromolecules, leading to the total understanding of the organisms made up of these substances. In this way, biology has changed its primary approach from toxology to a more analytical approach involving chemistry or physics, caused partly by X-ray crystallography.

1.0.1 X-ray crystallography

The first successes in solving the tertiary structure of hemoglobin by J. Kendrew, and that of myoglobin by M. Perutz are also landmarks in molecular biology. These works proved that X-ray crystallography can reveal structures of molecules as large as proteins.

Limitations of X-ray crystallography

In spite of early successes, little progress in the subject had been made by the early 70's[1]. This was because X-ray crystallography of macromolecules has severe inherent difficulties. They are;

- Difficulty in crystallizing the macromolecule.
- Difficulty in data collection caused by the large unit cell and radiation damage.
- Problems in obtaining the phases of structure factor amplitudes.

At first, the macromolecules must be crystallized to diffract enough for structural determination. It often takes a few years for researchers to obtain an adequate crystal. Generally speaking, crystallization has been more of an art or technique than a science for the majority of crystallographers in the field of biology. This is partly caused by a lack of basic knowledge about the processes of crystallization. The best conditions for crystal growth need an optimal combination of such factors as protein concentration, temperature, pH, ionic strength, specific additives and precipitants [3]. Crystallizing macromolecules has been a series of 'trial and error' processes with much effort. This also means that X-ray crystallography greatly depends on chance.

Besides the difficulties of crystallization, crystallographers face further problems in data collection. The macromolecule crystals mainly consist of light atoms such as hydrogen, nitrogen, carbon and oxygen. They also contain a large amount of solvent up to fifty percent volume. These imply that the diffraction intensities become weak, especially at high resolution, and that signal to noise ratios tend to be worse due to the increasing noise level from the solvent.

Furthermore, the crystals usually have large unit cell dimensions of some tens to hundreds of angstrom, causing the spatial separation of diffractions to become severe. A large unit cell volume (V) also weakens the diffraction intensities, as shown by the following equation;

$$\overline{I(h\ k\ l)} \propto \frac{\overline{|F(h\ k\ l)|^2}}{V^2}$$

where $\overline{I(h\ k\ l)}$ is the intensity of a diffraction, $\overline{|F(h\ k\ l)|}$ the amplitude of a structure factor and V the unit cell volume [3]. The weaker diffraction intensities in macromolecular crystallography make a high resolutional analysis much more difficult than those of low molecular weight substances.

The solvents contained in the crystal cause a more serious problem—radiation damage. It is inferred that a solvated electron (e^-_{aq}), hydrogen

[1]Only three structures had been solved by 1965, and 11 by 1970.

atom(H) and hydroxyl radical (OH) are produced during exposure in the X-ray beam, followed by the formation of free radicals ($\dot{R}$) [4].

$$RX + e^-_{aq} \longrightarrow \dot{R} = \overline{X}$$

The solvated electron may also react with H^+,

$$H^+ + e^-_{aq} \longrightarrow H$$

In this way, the solvated electron reacts with the covalent bonds of the macromolecules and destroys the crystalline order in the X-ray, leading to the decay of diffraction intensities. Due to the radiation damage, it has been almost impossible to collect a whole dataset with one crystal. The dataset obtained by merging those of a few or more crystals sometimes gives an electron density map which is difficult to interpret.

The phase problem in X-ray crystallography

The phase problem seems to be one of the most serious difficulties in the X-ray crystallography for macromolecules. The intensity of the diffracted reflection $\mathbf{I}(h\ k\ l)$ is proportional to the square of the amplitude of the structure factor $\mathbf{F}(h\ k\ l)$. The structure factor is a function of the electron density distribution $\rho(x\ y\ z)$ in the unit cell,

$$\mathbf{F(S)} = \sum_j f_j exp[2\pi i \mathbf{r_j} \cdot \mathbf{S}]$$

where $\mathbf{S} = \mathbf{s} - \mathbf{s_0}$ ($|\mathbf{s}| = |\mathbf{s_0}| = 1$), s is a vector of a diffracted reflection and s_0, of the insident X-ray. The summation is over all atoms j in the unit cell. Instead of summing over all separate atoms we can integrate over all electrons in the unit cell:

$$\mathbf{F(S)} = \int_{cell} \rho(\mathbf{r}) \mathbf{exp}[2\pi \mathbf{i r} \cdot \mathbf{S}] \mathrm{dv}$$

where $\rho(\mathbf{r})$ is the electron density at position r in the unit cell. Suppose x, y and z are fractional coordinates in the unit cell ($0 \leq \mathrm{x} < 1$; the same for y and z) and V is the volume of the unit cell, we have

$$dv = V \cdot dx dy dz$$

$$\mathbf{r} \cdot \mathbf{S} = (a \cdot x + b \cdot y + c \cdot z) \cdot \mathbf{S} = a \cdot \mathbf{S} \cdot x + b \cdot \mathbf{S} \cdot y + c \cdot \mathbf{S} \cdot z = hx + ky + lz$$

Therefore, $\mathbf{F(S)}$ can also be written as $\mathbf{F}(h\ k\ l)$.

$\mathbf{F(h\ k\ l)} = \mathbf{V} \int_{\mathbf{x=0}}^{\mathbf{1}} \int_{\mathbf{y=0}}^{\mathbf{1}} \int_{\mathbf{z=0}}^{\mathbf{1}} \rho(\mathrm{x\ y\ z}) \mathrm{exp}[2\pi \mathbf{i}(\mathbf{hx + ky + lz})] \mathbf{dxdydz}$ Then, by Fourier transformation, we can calculate the electron density ρ at every

position x, y, z in the unit cell:

$$\begin{aligned}\rho(x\ y\ z) &= \frac{1}{V}\Sigma_h\Sigma_k\Sigma_l \mathbf{F}(h\ k\ l)\exp[-2\pi i(hx+ky+lz)] \\ &= \frac{1}{V}\Sigma_h\Sigma_k\Sigma_l |\mathbf{F}(h\ k\ l)|\exp[-2\pi i(hx+ky+lz)+i\alpha(h\ k\ l)]\end{aligned}$$

However, there is a problem. Although the $|\mathbf{F}(h\ k\ l)|$s can be derived from the intensities $\mathbf{I}(h\ k\ l)$, the phase angles $\alpha(h\ k\ l)$ can not be obtained directly from the diffraction pattern. This problem was an obstacle of the early X-ray crystallography, and will continue to be in future.

Crystallization techniques

Because of the limitations described above, it was very difficult to solve the structures of biological macromolecules by X-ray crystallographic method. The situation has largely changed in recent years. Besides progress in the fundamental understanding of crystallization, the discovery and use of effective additives has contributed largely to success of X-ray crystallography. For instance, the use of polyethylene glycol (PEG) as an additive has accelerated the rate of success [5]. PEG has many advantages over salts such as ammonium sulphate, phosphate and citrate. It is effective at relatively low ionic strength, causing that the back ground level diffracted by a crystal from the PEG solution to be lower than those from higher ionic strength. For the researchers, this leads to easier interpretation of electron density maps of proteins. Furthermore, PEG has an additional advantage. The high effectiveness of low ionic strength offers higher ligand binding affinities to proteins than does a high ionic strength medium such as concentrated salt. This is highly significant in preparing isomorphous derivatives and complexes of macromolecules with substrate analogues, inhibitors and other macromolecules. It can be said that having such an effective precipitant as PEG has not only made the crystallization more successful, but also made the further processes of structural analysis easier. Nowadays, nearly half of the protein structures solved have used PEG [3], and it takes the prime position as the precipitating agent.

An empirical screening method has also accelerated the rate of success. In this method, a few sets of the most probable conditions for crystallization are selected based on successful examples of crystallization. [2] This systematic protocol has saved much exertion, and the possibility for crystallization has become much greater than that of previous decades.

[2] You can purchase the crystallization kit, CRYSTAL SCREEN™.

Progress in data collection

The average time required for data collection has been shortened owing to newly developed apparatuses in the X-ray field, enabling us to reduce radiation damage of crystals. The cryo-technique has been also developed in recent years, and has proved to be very useful in reducing X-ray damage during data collection [6, 7].

Development of a rotating anode generator has increased the X-ray intensity available in home-laboratories. The rotating anode generator can produce a more brilliant[3] and fine-focused X-ray beam. The development of two-dimensional detectors such as the image plate, the multiwire area detector and the charge coupled device (CCD) also has much contributed to reduce the measurement times.

Although the X-ray beam from the rotating anode generator is high brilliant, that from synchrotron radiation is two orders of magnitude stronger than the a conventional X-ray. It is especially suitable for weakly diffracting specimens, such as very tiny crystals or crystals with extremely large unit cells. Moreover, the beam has a low divergence resulting in sharper diffraction spots and high signal to noise ratio[4]. Its tunability in the wavelength is beneficial when the anomalous dispersion effects are used for MAD (Multiple wavelength Anomalous Dispersion) measurement.

Phase improvement

The phase problem has also been greatly improved by the development of many useful algorithms. In calculating phases of reflections by multiple isomorphous replacement (MIR), it is often very difficult to calculate the phases precisely with only two isomorphous derivatives, even though this is possible in theory. The MIR phases generally have lack-of-closure errors, causing much noise in the initial electron density map. The errors in phases can be effectively improved by recently developed techniques; solvent flattening [8], molecular averaging, non-crystallographic symmetry averaging and density modification [9].

As mentioned above, in the early stage of the subject, determination of the macromolecule structures was extremely difficult due to several obstacles and strongly depended on chance. Therefore, crystallographers tended to select areas of study depending on how readily they could obtain and crystallize the materials not regarding how significant the specimens were in biology.

However, the stance of crystallographers on X-ray analysis has been changing in recent years. X-ray crystallography of macromolecules has great-

[3]Brilliance is defined as number of photons/sec/mrad^{2}mm^2/0.1% relative bandwidth.

[4]In the present analysis, the synchrotron radiation source in KEK Tsukuba, Japan was used because our crystal has large unit cell dimensions.

ly progressed in many aspects including crystallization, data collection and calculation of phases. Nowadays the determination of three-dimensional structures of macromolecules is not as difficult as it used to be. Crystallographers have come to positively choose the objects of study with respect to their importance in biology and for further understanding of their functions. For instance, the crystal structure of HIV acid proteinase was solved for the purpose of developing specific inhibitors to prevent virus proliferation leading to the development of AZT [10]. A specific inhibitor of angiotensin-converting enzyme was developed based on its structure revealed by the X-ray technique, and is extremely useful in the treatment of hypertension or heart diseases [11]. In recent biology, it has been essential to know tertiary structures of biological macromolecules at atomic resolution. Nowadays, X-ray crystallography seems to have reached a stage where everyone without much knowledge of the method can study the structures they are interested in.

Chapter 2

Native structure

2.1 Introduction

The author has been interested in the relationship between the structure and function of biological substances, especially of enzymes. He has selected the X-ray crystallographic technique for exploring this relationship, because it has a powerful ability to offer him the finest information on tertiary structures of biological macromolecules at atomic resolution. He decided to carry out the X-ray crystallographic analysis of an enzyme, termed alkaline protease, from *Pseudomonas aeruginosa* IFO3080. The alkaline protease is a kind of bacterial zinc metalloendoproteases[1] whose tertiary structure was unknown. This enzyme consists of 470 amino acid residues. Its optimum pH is in slightly basic range, 7–9 [12, 13], as summarized in Table 2.1. The specificity of the enzyme is reported to be relatively broad toward natural and synthetic substrates, which contrasts with the specificities of thermolysins or carboxypeptidases. The enzyme is classified as one of metalloproteases, and protease possesses one zinc ion in the active site, which plays crucial roles in the substrate binding and increasing nucleophilicity of a water molecule bound to the zinc ion.

Significant roles of zinc for enzymes

Zinc is an essential component of many enzymes involved in various aspects of biological processes such as proteolysis and gene expression. It has been found to be an indispensable component for nearly three hundred enzymes in different specieses, including zinc proteases. The metal plays crucial roles, not only in catalysis (*c.f.* Figure 2.2) but also in stabilization of proteins and nucleic acids as structural zinc.

For instance, asparate transcarbamoylase (ATCase) contains one struc-

[1]More accurately, the alkaline protease should be termed alkaline 'proteinase'.

Table 2.1: General properties of alkaline protease. The number the calcium ions was determined through this X-ray analysis.

Strain	*Pseudomonas aeruginosa* IFO3080
Superfamily	metzincins
Family	serralysins
Molecular weight	≅50,000
No. of residues	470
Optimum pH	7–9
Optimum temperature (°C)	60
Ligands	1 Zn^{2+}
	8 calcium ions

tural zinc atom with two cystein residues as zinc ligands. The zinc atom contributes to the quaternary structure stabilization [14]. Zinc finger motif widely occurs in DNA binding proteins, where the metal serves as the structural zinc, typically liganded by two histidine and two cystein residues [15]. In an alcohol dehydrogenase (ADH)[2] [16], one zinc atom has a structural role, and the other a catalytic role. The atom can have either a structural or a catalytic role. The structural zinc atom is ligated tetrahedrally by the four sulfur atoms of four cystein residues.

Advantages of zinc for organisms

Why does nature select zinc as an integral component for such a wide range of organisms ? Zinc has some advantageous properties compared with other divalent metals such as Ni, Co, Cu and Hg. In focusing on the ionic radii of these metals, an increase is seen in the series [2];

$$\text{Ni(II)}\ (0.69\ \text{Å}) < \text{Co(II)}\ (0.72\ \text{Å}) < \text{Cu(II)}\ (0.72\ \text{Å}) < \text{Hg(II)}\ (1.1\ \text{Å})$$

Because of its compact size, zinc is almost as strong a Lewis acid as Cu(II) [17], which is suitable for polarizing other molecules (*c.f.* Figure 2.2, page 12. However, in contrast to Cu(II) and Co(II), zinc is inert to redox reactions, which may be an advantage in biological systems. In addition, it is relatively flexible in its coordination geometry [18].

[2]Handbook of Chemistry and Physics (1974–1975)

Zinc metalloproteases

Proteases which use zinc as a catalytic integral metal are classified into zinc metalloprotease. The superfamily is further subdivided into two types: zinc metallo-endo- and -exo-proteases. The representative enzymes of zinc metalloexoproteases are carboxypeptidases, whose catalysis mechanisms are among those best characterized through the native and complex crystal structures [19, 20, 21, 22, 23, 24, 25, 26]. Through sequence comparison among calboxypeptidases, the enzymes are found to conserve four ligands to the catalytic zinc atom; two histidines and one glutamic acid are found to be conserved in these enzymes. The forth ligand is also known to be a water molecules. Very few changes occur in the amino acid residues adjacent to them [27, 28, 29, 30, 31, 32, 33, 34, 35, 36]. The exo-catalytic manner of the enzymes is attributed to the S_1' pocket which is essentially a deep dead-end pocket for the terminal P_1' residue, and the structure is unable to accommodate any other residues at the further primed positions[3]. The zinc metalloendoproteases termed 'zincins' are hitherto classified into five distinct superfamilies as shown in Figure 2.1, termed 'zincins' [38, 39]. Until 1993, no tertiary structures from the a superfamily termed 'metzincins' had been available. The further classification which subdivides the zincins into two superfamilies was carried out by W. Bode *et. al.* after strong similarities were found between the crystal structures, including the zinc binding sequences and the 'met-turn' structure of astacin [40] and adamalysin [41]. The zincins have a conserved motif for the zinc binding site, **HEXXH**,[4] in an 'active-site helix'. The thermolysins, which are one of the enzyme most intensely studied by X-ray crystallography, have four zinc ligands, two histidines in the motif, and one glutamic acid in a position about twenty residues downstream from the second histidine ligand, and a catalytically important water molecule.

In contrast, the other four families can be further classified into a superfamily 'metzincins', which exhibit the extended motif, **HEXXHXXGXXH**. The three histidines in the consensus motif were found to serve as zinc ligands, whose manner of zinc ligation is different from that of thermolysins, and the glutamic acid is inferred, based on the knowledge of thermolysins [42], to act as a 'general base' residue to polarize a water molecule involved in a nucleophilic attack at the scissile peptide bond. Thermolysin, which is one of the bacterial zinc metalloendoproteases, has offered us a knowledge of tertiary structure solved by X-ray crystallography [43], and this structure has been the standard reference for those of other metalloproteases. The catalytic mechanism of thermolysin was proposed based on the native structures and the complexes with inhibitors and transition state analogues

[3]The nomenclature used for the individual amino acid residues R_2, R_1, R_1', R_2', etc. and the active site subsites S_1, S_1', etc. follows that of Schechter & Berger [37].

[4]X is a variable amino acid residue.

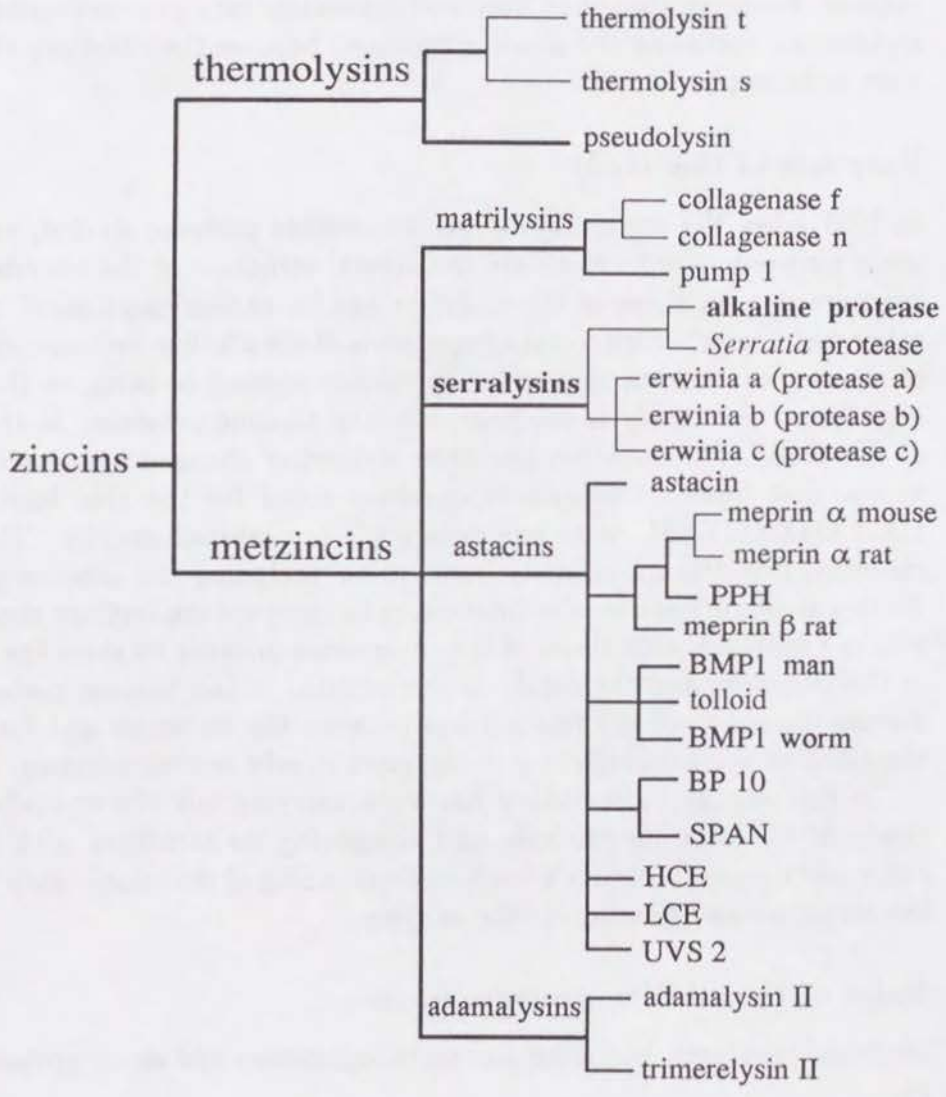


Figure 2.1: Five zinc metalloendoproteases. Zinc metalloendoproteases are classified into two superfamilies: thermolysins and metzincins. In each enzyme of the metzincins, a consensus methionine is conserved at the Met-turn. The metzincins are further subdivided into four families: matrixins, serralysins, astacins and adamalysins. Only representative members of the metzincins are shown.

[44, 45, 46, 47, 42, 48]. As shown in Figure 2.2, the catalytic zinc ion liganded to the active site cleft plays a crucial role in the binding and depolarizing of the carbonyl oxygen of the substrate, leading to the hydrolysis of the scissile peptide bond [47, 49, 42], where Glu143 is postulated as a general base residue. However, there had been no proposal for catalytic mechanisms of the metzincins, including the alkaline protease, because their tertiary structures were unknown until 1993.

Purposes of this study

In 1991 when the present study on the alkaline protease started, one of the main purposes was to elusidate the crystal structure of the enzyme and to compare it with those of thermolysins and/or carboxypeptidases, not with other metzincins[5]. Structural comparisons of the alkaline protease with those of structurally known zinc metalloproteases seemed to bring us the knowledge about the catalytic mechanism of the alkaline protease. In the course of the study, the situation has been somewhat changed because some enzymes that have an extended consensus motif for the zinc binding site, HEXXHXXGXXH, were analyzed by X-ray crystallography. They were classified into the superfamily 'metzincins' including the alkaline protease. At this stage, it became also interesting to compare the tertiary structure of alkaline protease with those of the metzincins in order to show the features of the structure and the catalytic mechanism. It had become customary to discuss questions of the relationship between the structure and function of the alkaline protease referring to the more closely related enzymes.

In this context, the author has been carrying out the crystallographic study of the alkaline protease and comparing its structure with those of other metzincins to obtain a basic understanding of the relationship between the structure and function of the enzyme.

Roles of the alkaline protease in vivo

Bacterial proteases, including zinc metalloproteases and serine proteases such as,

- *Pseudomonas aeruginosa* alkaline protease and elastase.
- *Serratia marcescens* 56-, 60- and 73-kDa proteases.
- *Aspergillus melleus* protease.
- *vibrio vulnificus* protease, subtilisin from *Bacillus subtilis*.
- Thermolysin from *Bacillus* stearothermophiilus.

[5]There was no classification, 'metzincins' at that time.

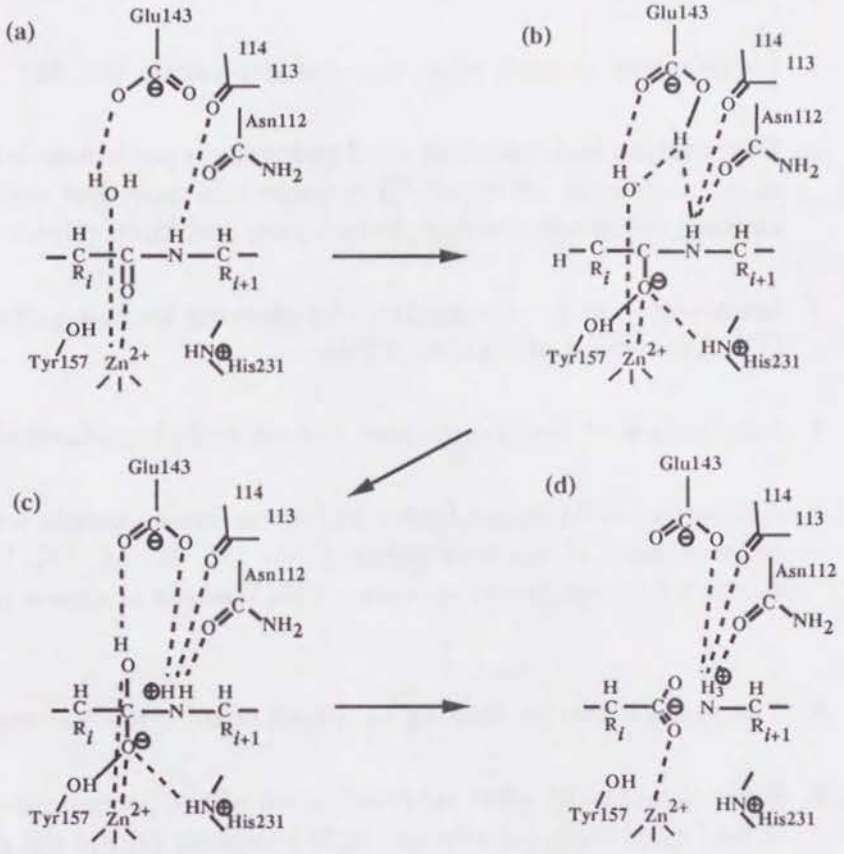


Figure 2.2: Proposed catalytic mechanism of Thermolysin. (a) The carbonyl oxygen atom of a bound peptide substrate ligands to the catalytic Zn^{2+} ion, and the water molecule must be pushed toward Glu143 carboxyl. (b) The pushed water molecule is increased in its nucleophilicity by the Zn^{2+} and the Glu143. It reacts with a carbonyl carbon of the scissile peptide bond and (c) generates the tetrahedral intermediate or transition state, which is further stabilized by electrostatic interactions of the carbonyl oxygen with the Zn^{2+} and by two hydrogen bonds from other side chains. The proton has been tranferred from the Glu143 to -HN- of the scissile peptide bond. (d) The tetrahedral species leaves a proton on the Glu143 carboxyl, then it moves onto -NH- of the scissile bond to cleave the peptide bond.

- *Streptomyces* caespitosus protease and V8 protease from *Staphylococcus* aureus.

cause a vast and complex range of pathogenic effects in infected foci. These effects are;

1. Liquefactive necrosis when injected into cornea [50, 51].

2. Degradation and inactivation of endogenous proteinase inhibitors, such as α_1-proteinase inhibitor, C1-esterase inhibitor, and antithrombin III, causing unlimited clotting, fibrinolysis, and kinin release [52, 53].

3. Inactivation of immunosystems by cleaving immunoglobulin G and A [53], lysozyme and transferin [54].

4. Inactivation of the complement system and chemotaxis [55].

5. Activation of Hageman factor–kallikrein–kinin cascade resulting in the enhancement of vascular permeability [51, 55, 56, 57], leading to activation of coagulation cascade. This cascade is shown in Figure 2.3.

6. Tissue degeneration leading to spread of infection and septicemia [50].

7. Sever cytotoxicity after internalization of the proteinase–α_2 macroglobulin (α_2M) complex through α_2M receptors on the cell surface [58].

8. Activation of factor X and plasminogen, causing the increase of influenza and paramyxovirus infectivity [59].

9. Activation of bacterial toxin, leading to enhanced spreading of bacterial toxins mediated through kinin activation [59].

The Hageman factor, in the form of a zymogen, is a kind of serine proteases, which is activated by many kinds of bacterial proteases. The Hageman factor dependent pathway is considered to be an initiation system of the intrinsic blood coagulation pathway. The pathway has three zymogen-type proteases (Hageman factor, prekallikrein and blood coagulation factor XI) and one cofactor (high-molecular-weight kininogen). Activation of these zymogens is associated with an amplification mechanism in the protease cascade system [55].

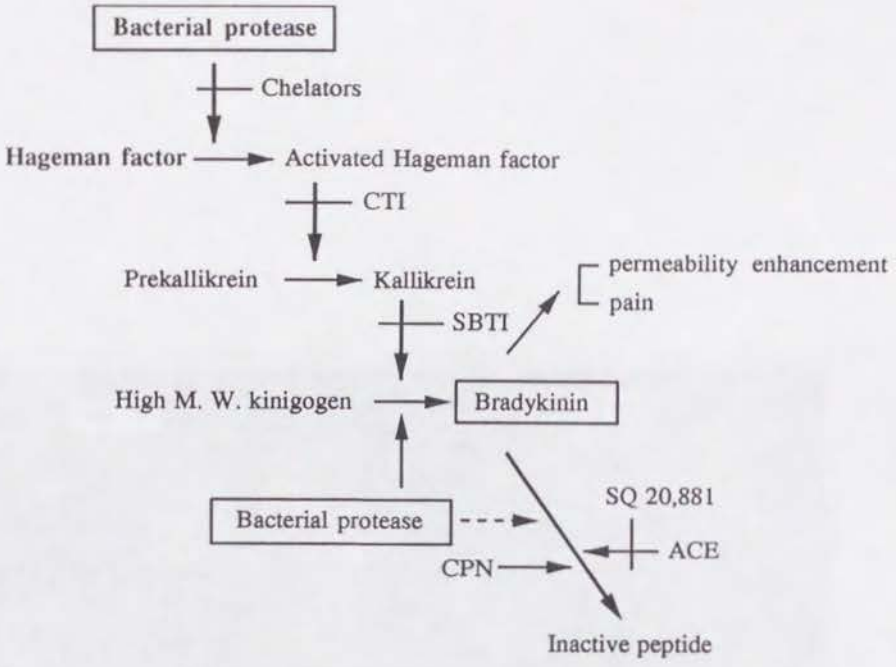


Figure 2.3: Schematic representation of the activation of the Hageman factor-kallikrein-kinin cascade by bacterial proteases such as *Pseudomonas aeruginosa* alkaline protease or *Serratia* protease. CTI; corn trypsin inhibitor, SQ 20,881; EYPRPQIPP-OH, ACE; angiotensin converting enzyme, CPN; carboxypeptidase N [55].

Pseudomonas aeruginosa

The bacterium *Pseudomonas aeruginosa* is a species of gram-negative, aerobic, rod-shaped bacteria commonly isolated from clinical specimens (wound, burn, and urinary tract infections). It is also widely distributed in soil and water. The bacterium is a major agent of nosocomial and infection a major causative agent of opportunistic infection. It is a kind of gram-negative bacteria and secretes a large amount of bacterial toxins into extracellular mediums[6], which contrasts with the way that normal gram-negative bacteria secrete the toxins because many of them produce endotoxins.

The alkaline protease was first isolated by Dr. K. Morihara [60, 13], and inferred as one of the exotoxins secreted by *Pseudomonas aeruginosa* strains IFO3455, T–30, IFO3080 and P–M, as shown in Figure 2.4. The strains IFO3455 and T–30 further secrete *Pseudomonas aeruginosa* elastases [61], which were classified as thermolysins (*c.f.* Figure 2.1, page 10) and

[6]exotoxins

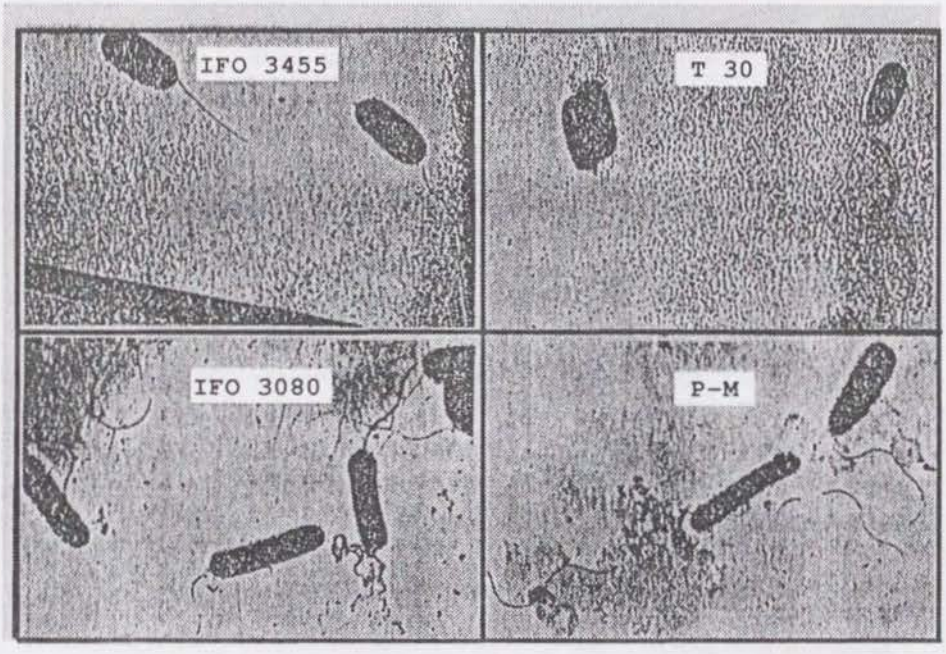


Figure 2.4: Electron microscopic photographs of *Pseudomonas aeruginosa* strain IFO 3455, T–30, IFO3080 and P–M. The strains IFO 3455 and T–30 produce *Pseudomonas aeruginosa* elastases, but neither those of IFO3080 nor P–M does. These photographs were kindly provided by Dr. K. Morihara.

their tertiary structure is found to be similar to that of thermolysin [62]. Those strains with the elastase have been shown to be much more pathogenic than those without it; their toxity was partly shown by an acute liquefactive necrosis in the guinea pig cornea and permeability enhancement reaction of a vascular in guinea pig skin caused by these enzymes [50].

Therefore it has been suggested that the elastase and the alkaline protease are, in cooperation with each other, probably responsible for the tissue destruction observed during pulmonary and corneal infections by this pathogen [63].

2.1.1 Specificity of the alkaline protease

In an early study on characterization of the alkaline protease, K. Morihara *et al.* reported that the enzyme exhibited a broad specificity toward the oxidized insulinB-chain [12], as shown in Figure 2.5. In a later study on the

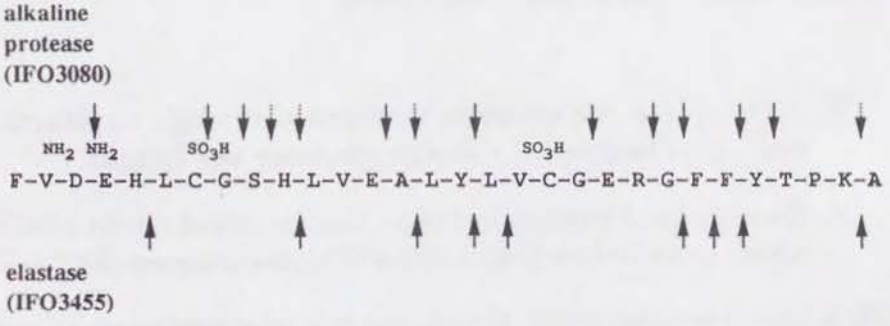


Figure 2.5: Broad specificity of the alkaline protease demonstrated by hydrolysis of an oxidized insulin B-chain.

specificity of the enzyme, however, a result inconsistent with the previous one was reported [64, 65, 66]. Synthetic peptides were systematically hydrolyzed by the alkaline protease as shown in Table 2.6. This experiment led to the conclusion that the enzyme had the major recognition site at P_1 for Arg (or Lys) with an additional site at P_1' because the the residue enhanced susceptibility of the enzyme to the substrates. It is essential to solve the crystal structure of the alkaline protease to discuss the discrepancy beteween the results from these experiments. In summary, the main purposes of this study are;

1. To solve the crystal structure of the alkaline protease in order to obtain a fundamental knowledge of relationships between the structure and function of the enzyme.

Synthetic substrate	Km (mM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
$P_4 - P_3 - P_2 - P_1 - P'_1 - P'_2 - P'_3$			
bz-Gly↓Phe-Arg↓Ala-Nba	0.20	0.11	0.51
Abz-Gly↓Phe-Arg↓Val-Nba	0.25	0.03	0.12
Abz-Gly↓Phe-Arg↓Leu-Nba			
Abz-Gly↓Phe-Arg↓Phe-Nba	0.13	0.13	1.10
Abz-Gly↓Phe-Lys↓Leu-Nba	0.23	0.03	0.15
Abz-Gly-Phe-Arg↓Leu-Gly-Nba	0.25	0.29	1.30
Abz-Gly-Phe-Arg↓Leu-Leu-Nba	0.04	0.32	7.50
Abz-Met-Thr-Arg↓Leu-Leu-Nba	0.16	0.73	4.50
Z-Phe-Arg↓MCA	0.14	0.0027	0.02
Abz-Ala-Gly-Leu↓Ala-Nba	0.13	0.12	0.98

Figure 2.6: Kinetic parameters on the hydrolysis of fluorogenic substrates by the alkaline protease. Arrows indicate the bonds to be split. The experimental conditions were as follows; 50 m*M* Tris buffer, pH8.0/0.1 *M* NaCl, 0.01 M $CaCl_2$, 2.3 % DMF, and 25°C. Abz-; amino benzoyl group, -Nba; 4-nitrobenzyl group. Dr. Nishino *et. al.* concluded that the hydrolysis of -Gly-Phe- bonds was attributed to the contamination of *Pseudomonas aeruginosa* elastase into the specimem.

2. To elucidate the catalytic mechanism through structural analysis of complexes between the alkaline protease and ligands.

3. To compare the structures of metzincins including the alkaline protease, which reveals their functional differences induceed by the structures.

To achieve these purposes, the author has intensively carried out the structural analysis of the alkaline protease and its complexes by X-ray crystallography at high resolution.

2.2 Materials and Methods

2.2.1 Crystallization

Although X-ray crystallography has made great progress as described in Chapter 1, one point remains unchanged— success in solving the structures largely depends on the quality of crystals. To obtain crystals of the alkaline protease suitable for X-ray analysis, the following experiments were conducted. The lyophilized sample of the alkaline protease from *Pseudomonas aeruginosa* IFO3080 (Nagase Biochemical Co. Ltd), was kindly gifted by Dr. K. Morihara[7]. The incubation and purification parameters of the specimem is summarized in Table 2.2. Crystallization trials were conducted using a hanging drop vapor diffusion method as illustrated in Figure 2.7. A large number of conditions were screened, varying the nature of precipitant, protein concentration, pH, additives and temperature[8]. The chemicals used and conditions tested are showed below.

Table 2.2: Incuvation and purification parameters for the alkaline protease.

Origin	*Pseudomonas aeruginosa* IFO3080
Quality	crystalline
Activity	4,400 units (mPU) per mg
Absorbance	O.D. (280 nm) of the solution of 0.2 mg/ml with distilled water; 0.259
Specific activity	5.43 mPU per mg of protein.*
Purity	108.7 %

*$E^{1\%}_{1cm}$ at 280 nm = 16.0

Hanging drop vapor diffusion method

Vapor diffusion techniques are most widely used for crystallization in macromolecular X-ray crystallography. Prior to a hanging drop vapor diffusion [5], 20 mm diameter glass cover slips (matsunami micro cover glass, matsunami glass ind., Japan) were washed with a silicon solution (Sigmacoat; Sigma chemical company, U. S. A.) and dried overnight at room temperature. The

[7]Institute for Applied Life Science, University of East Asia, Ichinomiya-Gakuen-cho 2-1, Shimonoseki, Yamaguchi Pref. Japan.

[8]All chemicals used were of extra reagent grade and purchased from Nacalai tesque,inc. Kyoto, Japan, without any notations. Water was singly distilled.

well rims of the micro plate (Cell Wells™; Corning Glass Works, Corning, New York, U. S. A.) were daubed with silicon compound (Toray Dow Corning Silicone Co. Ltd., Tokyo, Japan). Cystals suitable for X-ray analysis were prepared as shown in Figure 2.7. The protein solution was prepared by dissolving the 10 mg of the specimen in 500 μl of an 1 mM NaN_3 + 5mM $CaCl_2$ solution (pH 7.0). The wells were filled with a reservoir solution. The reservoir solution was 50 mM acetate butter (pH 5.6) containing 6 % (w/v) polyethylene glycol 6000 and 1mM NaN_3. Hemispherical 5 μl droplets of the protein solution were formed on the glass slips by mixing the protein solution and the reservoir solution in the volume ratio 1:1. The cover slips were overturned and set on the rims and pressed to seal the wells with the silicon compound. Small crystals (in the maximum size $\leq$ 0.3 mm) as shown

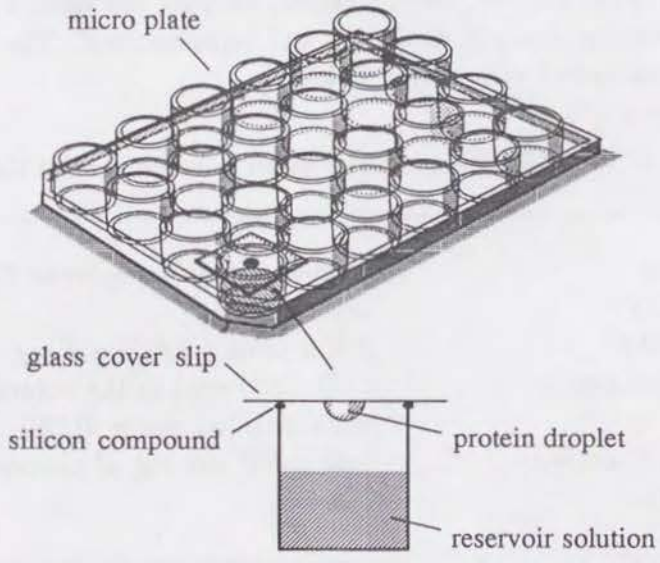


Figure 2.7: Hanging drop vapor diffusion method. The protein droplet consists of 2.5 μl of a 2 % (w/v) protein solution (5mM $CaCl_2$/1mM NaN_3, pH 7.0) and a 2.5 μl reservoir solution. The reservoir solution consists of 50 mM acetate buffer (pH 5.6) containing 6 % (w/v) polyethylene glycol 6000 and 1 mM NaN_3. Each well was filled with 1 ml of the reservoir solution. The micro plate was placed in an incubator at 25 °C.

in Figure 2.8 were obtained in the hanging drops.

Macroseeding technique

Blood coagulation glass plates as shown in Figure 2.9 (six depressions, 75×49-×5 mm) were washed with the silicon solution and then immediately rinsed

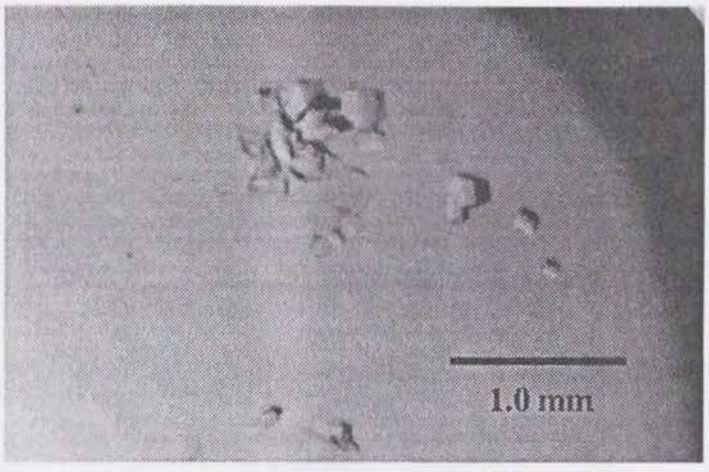


Figure 2.8: Small and stacked crystals. They were obtained from the hanging drops at 25 °C, typically in 3–7 days.

with water. The rim of a plastic box was daubed with the silicon compound. Both the protein solution and the reservoir solution were filtered through cellulose acetate filters (13 mm diameter 45 μm pore diameter, Disposable Sterile Syringe Filter^{TM}; Corning Glass Works, Corning, New York, U. S. A.) and used for further experiments. Small crystals (0.1-0.3 mm) obtained by the hanging drop vapor diffusion method, which must be free from twinning or any other crystalline state, were used for a macroseeding technique [67] to grow the crystals. The macroseeding technique was performed as follows;

1. A micro vessel was placed on each of the four corners of a plastic box. 1 ml of the reservoir solution was poured into each of the vessels.

2. One of the small prismatic crystals was picked up into a Pasteur pipet connected to a micro syringe with a piece of rubber tubing, under a dissecting microscope using a magnification of 10–40 $\times$.

3. To wash out the micro crystal, the prismatic crystal was introduced from the pipet into a the reservoir solution on a blood coagulation glass plate.

4. The washed crystal was again picked up into the pipet, and brought into a droplet consiting of 30 μl of the protein solution and the equal volume of the reservoir solution, on another blood coagulation glass plate.

5. Each of six cystals was set into the each droplet on the plate, and the plastic box was sealed with a transparent plastic cover, one by one.

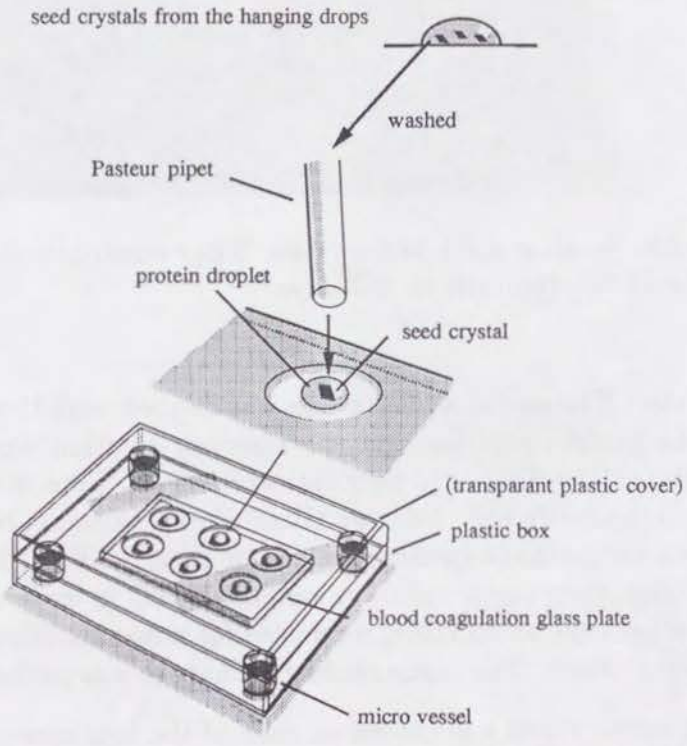


Figure 2.9: The macroseeding procedures. A seed crystal was picked into a droplet of the precipitant solution to wash out the micro crystals attaching to the seed. The seed was then introduced into 60 μl of the protein solution which consists of 30 μl of the 2 % solution and the equal volume of the reservoir solution. The rim of the plastic box was daubed with silicon compound. The box was sealed with a transparent plastic cover, and placed at 25 °C.

The procedures are illustrated in Figure 2.9. The small crystals obtained from the hanging drop method could be reproducibly grown up to a size (typically, 1.0×0.6×0.3 mm) large enough for X-ray experiments using the macroseeding technique. The crystals grown were free from any twisting and stacking. A typical crystal shape is shown in Figure 2.10.

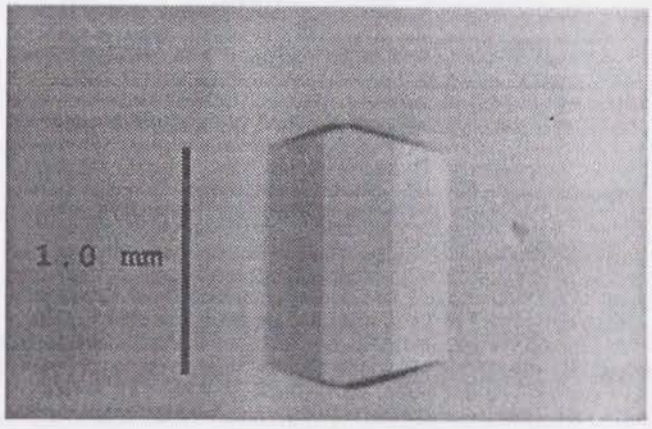


Figure 2.10: A prismatic crystal suitable for X-ray experiments. The crystal, grown by the macroseeding technique, typically reached to a size of 1.0×0.6×0.3 mm in 7–10 days.

Precession photograph

To determine the crystallographic parameters, and assess crystallinity of the target crystal, a photographic method is still widely used. Precession photographs were first taken for these purposes. Before taking precession photographs of the crystal, the following items were prepared;

1. Pasteur pipets were drawn out in a flame to give a narrow bore and some were cut off to increase the bore at the tip.

2. The tip of an X-ray capillary (1.5 mm diameter 80 mm long, Mark-Röhrchen für röntgenographische Aufnahmen, Berlin, Germany) was snapped open, and soaked into the silicon solution, then rinsed with water immediately, and dried. The open end of the capillary was heated to close it.

3. The upper half of the capillary was filled with the reservoir solution using the narrow Pasteur pipet.

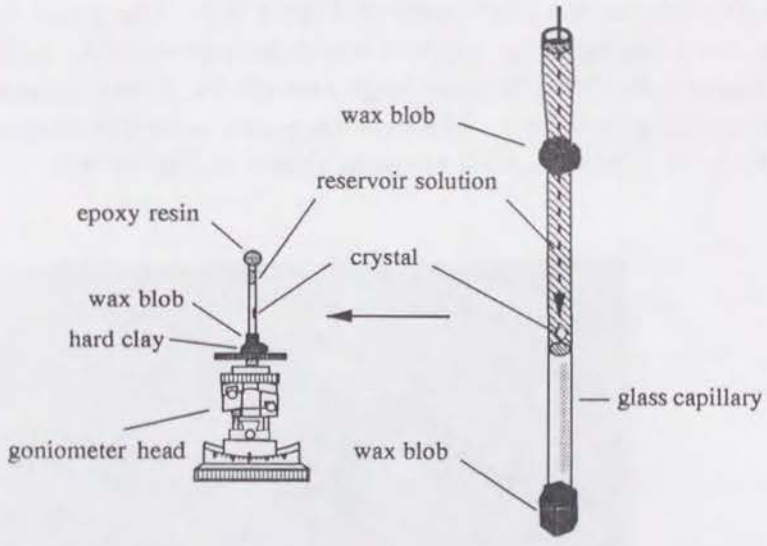


Figure 2.11: Crystal mounting method. The closed tip of the glass capillary was fixed to the wax blob (GC New Sticky Wax; GC Dental Products Inc. Ltd., Tokyo, Japan). Another wax blob was daubed on the capillary near the open side. The crystal, picked up with the narrow Pasteur pipet, was introduced into the glass capillary and guided to the middle with the reservoir solution. The excess solution was removed from the capillary, leaving only a small amount of the solution surrounding the crystal. The upper part of the upper wax blob was clipped, and then the capillary was sealed with a small amount of epoxy resin, and fixed to the X-ray goniometer head with hard clay.

As shown in Figure 2.11 on page 23, the prismatic crystal grown to 0.5-1.0 mm in its maximum dimension by the macroseeding method, was selected for the precession photographs [4]. The crystal in the droplet was picked up with a narrow Pasteur pipet connected to the micro-syringe with a piece of rubber tubing, and then transferred into the capillary.

When the crystal got to the middle of the capillary, it was adhere to the inside wall of the capillary by sucking the reservoir solution into the narrow Pasteur pipet. The direction of the crystal was adjusted using a piece of thin glass fiber, and the excess solution around the crystal was absorbed with a thin filter paper to adhere the crystal to the inside wall. A small plug of the reservoir solution was introduced into the capillary over the crystal using the Pasteur pipet to keep the inside humidity constant. The open end of the capillary was sealed with epoxy resin (Araldite; Ciba-Geigy, Switzerland).

Crystallographic parameters were determined from the precession photographs shown in Figure 2.12. All of the photographs show that the crystal diffracts

Table 2.3: Conditions for taking precession photographs.

X-ray film	Kodak, 12.7×12.7 mm
Precession camera	Nonius
X-ray generator	Rigaku RU-200
Focus size	0.3×3.0 mm, fine
Collimator diameter	0.5 mm
X-ray power	40 kV, 100 mA
Exposure time	12–16 hours
μ	12°
Camera distance	100 mm
Slit radius	15 mm
Slit width	1 mm
Temperature	room temperature
Beam direction	*a*, *b*, *c*

up to at least 3 Å resolution, and that the diffaction spots were fine enough for X-ray analysis.

Conditions for taking precession photographs

The precession photographs were taken at the Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan. The crystal mounted on a goniometer was set on a precession camera (Nonius). Three photographs were taken with each crystallographic axis parallel to the incident X-ray beam. Detailed conditions are shown in Table 2.3.

Refinement of crystallographic parameters

The cell dimensions and space group determined by the precession photographs were further refined and examined with a diffractometer (Rigaku AFC-7). Operating conditions are shown in Table 2.4. The number of molecules per unit cell (Z) can be estimated by a method proposed by Matthews [68]. In most protein crystals, the ratio (V_M) of the unit cell volume and the molecular weight (M_W) is between 1.7 and 3.5 $Å^3$/Da. Assuming that $M_W = 50{,}000$ and $Z = 4$, then

$$\begin{aligned} V_M &= V_m/M_W/Z \\ &= 698,431.5/50,000/4 \\ &= 3.49(Å^3/Da) \end{aligned}$$

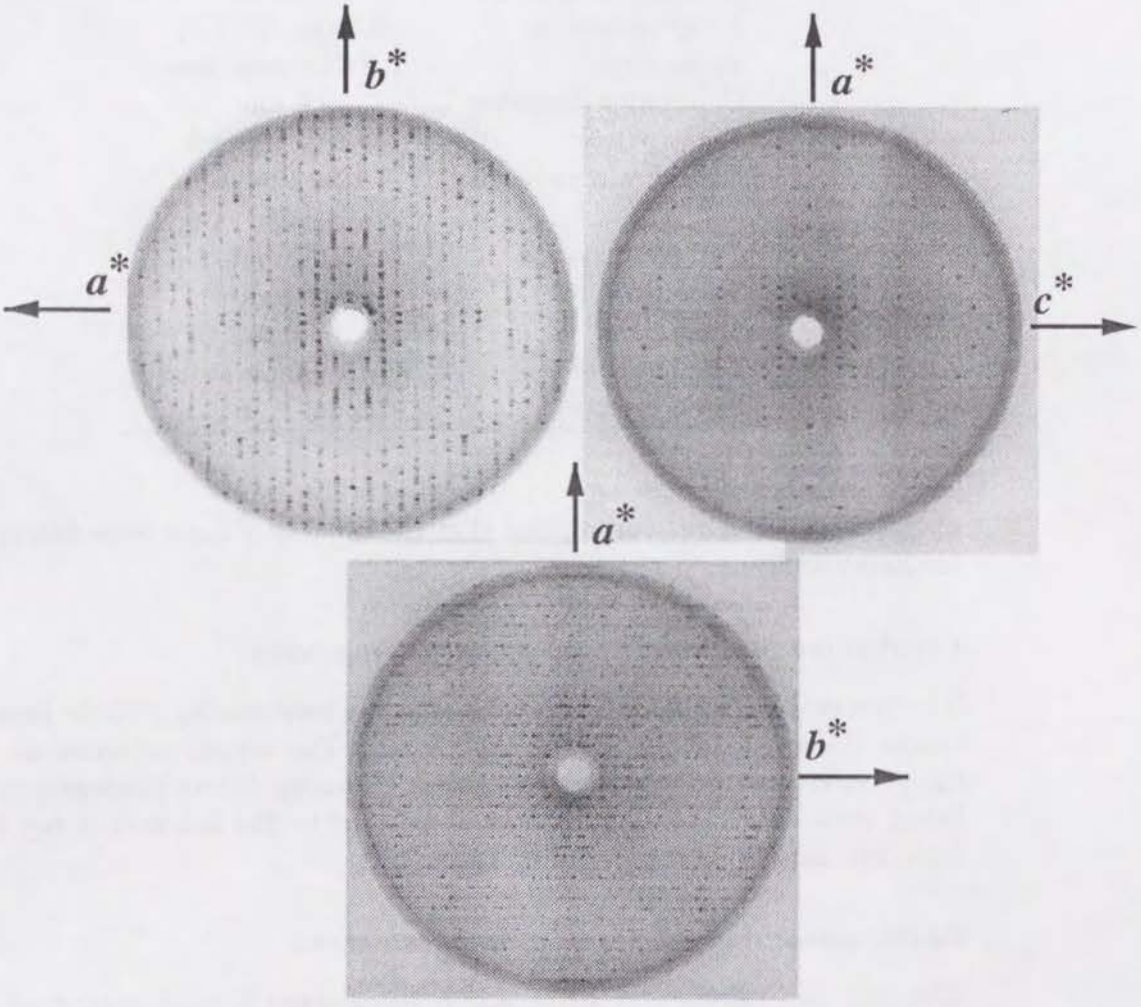


Figure 2.12: Precession photographs. These photographs indicate that the native crystal diffracts up to 3.0 Å resolution, and that the spots are fine enough for X-ray analysis.

Table 2.4: Operation conditions for diffractometer. This condition was further applied to derivative searches.

X-ray generator	RU-300
Focus size	0.5×10 mm, normal
Collimator diameter	0.5 mm
X-ray power	40 kV, 300 mA
Scan speed	5 deg./min.
Scan width	0.8 deg.
Measurement time for background	3.0 sec.
Slit size	0.5 deg.
Temperature	10.0 °C

There is an approximate relation between the volume of solvent (V_{solv}) in the unit cell ($V_{protein}$) and V_M:

$$V_{solv} = 1 - 1.23/V_M$$

This gives

$$\begin{aligned} V_{sovl} &= 1 - 1.23/3.49 \\ &= 0.65 \end{aligned}$$

Crystallographic parameters obtained are summarized in Table 2.5. Four molecules of the alkaline protease are packed in the unit cell as shown in Figure 2.13.

2.2.2 Preparation of heavy-atom derivatives

Protein crystals normally contain crystallization solvent by about 50 % of the crystal volume. Various kinds of chemical species such as ions, inhibitors, substrate analogues and heavy atoms can go through three channels and react with the protein without disturbing the crystalline order. These reactions can be achieved in full by soaking native crystals in ligand solutions [4]. When preparing isomorphous heavy-atom derivatives, following conditions must be satisfied;

- The heavy atoms must occupy a limited number of sites in the crystals, which are crystallographically identical.

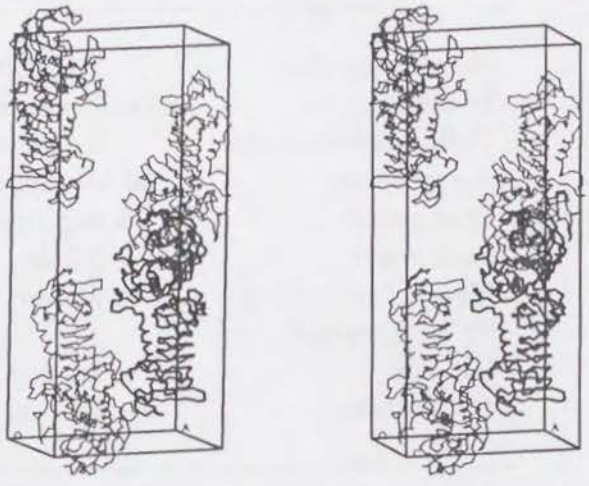


Figure 2.13: Crystal packing of the enzyme molecules in the unit cell. Each asymmetric unit contains one molecule. Note that there are solvent channels in the unit cell, which is significant for preparing the complexes.

Table 2.5: Crystallographic parameters of the native crystal. They were refined by the diffractometer.

Space group	$P2_12_12_1$
a (Å)	77.2
b (Å)	176.7
c (Å)	51.2
Z	4
V_m (Å^3)	693,431.5
V_M (Å)	3.49
V_{solv}	0.65

- The heavy atoms must react with the protein molecules without destroying the crystallinity.

Preparation of heavy-atom derivatives is basically a 'trial and error' process, just the crystallization.

Soaking of crystals

To solve the phase problem by multiple isomorphous replacement (MIR) [4], isomorphous heavy-atom derivatives were prepared by soaking method [4, 5]. Each heavy-atom agent was dissolved in the reservoir solution (*c.f.* page 19) to adjust the concentrations to 0.1–20 mM. The heavy-atom solution was poured into the upper half of the siliconized capillary fixed to the wax blob (*c.f.* Figure 2.11, page 23) with the narrow Pasteur pipet. Each crystal was soaked in the heavy-atom solution for a period of some hours to several weeks at 25 °C. The heavy-atom agents used and their concentrations are in Table 2.6. Prior to measuring diffraction intensities of the crystals soaked in the

Table 2.6: Soaking conditions. All of the heavy-atom agents were dissolved in the reservoir solutions (6 % (w/v) PEG 6000, 50 mM acetate buffer, pH 5.6, 5 mM $CaCl_2$/1 mM NaN_3).

Heavy-atom agent	Concentration (mM)	Soaking time (day)
$UO_2(NO_3)_2$	10.0, 2.5, 1.0, 0.5, 0.1	1–2
$K_3UO_2F_5$	10.0, 5,0, 1.0, 0.1	1–7
$UO_2(CH_3OCO)_2{\cdot}2H_2O$	0.1, 1.0	1
$Sm(NO_3)_3$	10.0, 5.0, 1.0, 0.2	1–3
$Er_2(SO_4)_3$	1.4	1–3
$CH_3COOHgC_5H_6$	*slightly soluble*	–
CH_3HgCl	15.0, 5.0	1–3
$HgCl_2$	2.5	1–3
K_2PtCl_4	*slightly soluble*	–
K_2PtCl_6	10.0	1–7
$Pb(NO_3)_2$	10.0, 1.0	1
CH_3OCOAg	1.0	1

heavy-atom solutions, intensity profiles on crystal axes ($h\ 0\ 0$), ($0\ k\ 0$) and ($0\ 0\ l$) were checked. The structure factor amplitudes of the crystals soaked in the heavy-atom solutions which seemed to give effective profile changes were

further measured to a resolution of 5.5 Å using the diffractometer under the same operating conditions as shown in Table 2.4, page 26.

Determination of the heavy-atom sites

The major heavy-atom binding sites were located using difference Patterson and difference Fourier maps calculated with the program package PHASES [69] installed on a workstation Iris Indigo (Silicon Graphics Inc., California, U. S. A.). Refinement of the heavy-atom parameters was performed by an iterative method of phase determination and least-squares refinement [70]. Minor heavy-atom sites of each derivative were found with difference Fourier maps during the refinement, and included in the further refinement.

Isomorphous heavy-atom derivatives

Three kinds of isomorphous heavy-atom derivatives, CH_3HgCl, $HgCl_2$ and $UO_2(CH_3OCO)_2{\cdot}2H_2O$ were obtained from the soaking experiments as shown in Table 2.6. The derivatives could be considered sufficiently isomorphous for the MIR method, giving enough changes in structure factor amplitudes compared with the native. Crystals soaked in other heavy-atom solutions such as $UO_2(NO_3)_2$, $Pb(NO_3)_2$, $(NO_3)_3$, and $Er_2(SO_4)_3$ were insufficient for the MIR method and lost their crystallinities. Those from $K_3UO_2F_5$ and K_2PtCl_6 gave slight changes in structure factor amplitudes.

Data collection using synchrotron radiation

Diffraction intensities of the native and derivative crystals were rather weak in high resolution range probably because of the high solvent content. Synchrotron radiation was required to collect the high resolution intensity data. All data sets for the crystals were collected with a screenless Weissenberg camera [71] at the $BL6A_2$ station of Photon Factory in National Laboratory for High Energy Physics, Tsukuba, Japan. The monochromatic X-ray wavelength was adjusted to 1.00 Å at the station. Image plates (Imaging Plate; Fuji Film Inc. Ltd., Tokyo, Japan) were used as two-dimensional detectors [72] for recording diffraction patterns. Image plates have many advantages over conventional X-ray films. Image plates are at least 10 times more sensitive than X-ray films and their dynamic range of the crystal is 10^4–10^5 times wider than that of X-ray films. In data collection using synchrotron radiation, their high sensitivity at the shorter wave length is especially useful.

Each of the two native crystals was mounted with its crystallographic b^*- or c^*-axis parallel to the rotation axis of the crystal, and all the derivatives were in the c^* -setting. For one Weissenberg shot, the oscillation range and the ratio of crystal rotation to cassete movement were 4.7° and $1.0°{\cdot}mm^{-1}$, respectively. The exposure time was 14.1 sec. A series of diffraction images

which covers a crystallographically independent area was recorded for the crystals. Twenty-five frames of the partial image for each crystal were scaled and merged together using the program package WEIS [73]. Measurement conditions applied to all of the cystals are summarized in Table 2.7.

Table 2.7: Conditions for data collection using synchrotron radiation.

Wavelength	1.00 Å
Spindle axis	b^*, c^* (native)
	c^* (derivatives)
Overlap	0.5 deg.
Oscillation angle $\Delta\omega$	4.7 deg.
Cassette movement ΔZ	4.7 mm
ω/Z	1.0 deg./mm
Collimator diameter	0.2 mm
Rotation speed	2.0 deg./sec.
Number of oscillations	6
Image plate size	20×40 cm
Temperature	10.0 °C

Data processing using the program WEIS

All diffraction intensity data collected using the synchrotron radiation X-ray source were processed to structure factor amplitudes with the program WEIS [73]. Two sets of the native data obtained in the b^* and c^* rotation axis settings, were merged into the native data set for increasing the completeness of data. The processing criteria for the native data were more strict than those for the derivatives. The statistics for data processing are shown in Table 2.8. The data processed were turned to be sufficient for X-ray analysis as indicated by the values of R_{merge} and completeness of the data.

Determination of heavy-atom parameters

To determine binding sites, temperature factors and occupancies of heavy atoms, difference Patterson functions of the derivatives were calculated using the following equations:

$$P(\mathrm{uvw}) = \frac{1}{\mathrm{V}} \sum_{hkl} |\Delta F(\mathrm{hkl})|^2 \cos[2\pi(\mathrm{hu} + \mathrm{kv} + \mathrm{lw})]$$

Table 2.8: Statistics for data processing. The native data were derived from two data sets, and were processed using more strict criteria than those of the derivatives.

	Crystals			
	Native	CH_3HgCl	$HgCl_2$	$UO_2(CH_3OCO)_2$
No. of crystals	2	1	1	1
Rotation axis	b^*, c^*	c^*	c^*	c^*
Resolution (Å)	2.0	2.0	2.0	2.0
Total observations	60,923	66,307	66,808	67,894
$1c(I > \sigma(I))$				
Unique reflections	44,151	32,610	32,053	32,801
Completeness of data (%)	96.8	71.5	70.2	71.9
[a] R_{merge} (%)	5.28	4.81	5.17	4.38

[a] R_{merge} (%)= $100\Sigma_i|\langle I_i\rangle - I_i|/|\Sigma i I_i$; $\langle I_i\rangle$ is the average of I_i over all observed symmetry equivalents.

$$|\Delta F(\mathrm{hkl})| = ||F_{\mathrm{PH}}(\mathrm{hkl})| - |F_{\mathrm{P}}(\mathrm{hkl})||$$

or in short,

$$P(\mathbf{u}) = \frac{1}{V}\sum_{S}|\Delta I(S)|\cos[2\pi\mathbf{u}\cdot S]$$

where u, v, and w are relative coordinates of the Patterson map space and $|\Delta F(\mathrm{h\ k\ l})|$ is the difference in structure factor amplitude between the native and derivative crystals. Note that the Patterson function does not contain any phase information. Therefore, the Patterson function can be directly calculated with structure factor amplitudes of the native and derivative crystals, which are derived from the X-ray measurements. Major heavy-atom binding sites of these derivatives were determined by calculating the Patterson function maps using the program package PHASES [69]. These maps clearly showed the peaks for the major heavy-atom binding sites, as shown in Figure 2.14–2.14. Furthermore, other heavy-atom sites were interpreted one by one using difference Fourier maps based on the major heavy-atom sites found in the difference Patterson maps. The heavy-atom sites were refined using the program PHASES [69]. The conditions for preparing isomorphous derivatives are summarized in Table 2.9, and the heavy-atom binding sites are shown in Table 2.10. A mini-map of native electron density [3] without

Table 2.9: Conditions for preparing isomorphous heavy-atom derivatives. The heavy-atom reagents were dissolved in the reservoir solution.

	CH_3HgCl	$HgCl_2$	$UO_2(CH_3OCO)_2$
Concentration of heavy-atom agent (mM)	15.0	2.5	1.0
Soaking time (day)	3	3	1
Temperature (°C)	25.0	25.0	25.0

Table 2.10: Heavy atom parameters.

Derivative	Site number	X	Y	Z	B	G
CH_3HgCl	1	0.5106	0.3976	0.4253	31.69	1.259
	2	0.0167	0.3993	0.1725	57.88	0.465
	3	0.0980	0.2425	0.5064	80.00	0.513
	4	0.7480	0.0837	0.3228	63.12	0.513
$HgCl_2$	1	0.0709	0.3240	0.2686	47.37	1.180
	2	0.0100	0.3996	0.1554	78.21	0.799
	3	0.0893	0.2488	0.4730	68.84	1.039
	4	0.5082	0.3973	0.4219	30.34	0.856
	5	0.7217	0.4529	0.4304	54.08	1.086
$UO_2(CH_3OCO)_2$	1	0.1877	0.1078	0.2686	49.87	0.684
	2	0.6800	0.2239	0.4883	45.66	1.486
	3	0.5633	0.1655	0.4448	43.60	0.387
	4	0.3305	0.0089	0.1091	55.22	0.414
	5	0.3290	0.2940	0.1562	56.78	0.311
	6	0.5076	0.4500	0.5079	67.92	0.345

B: temperature factor. G: occupancy in arbitrary scale.

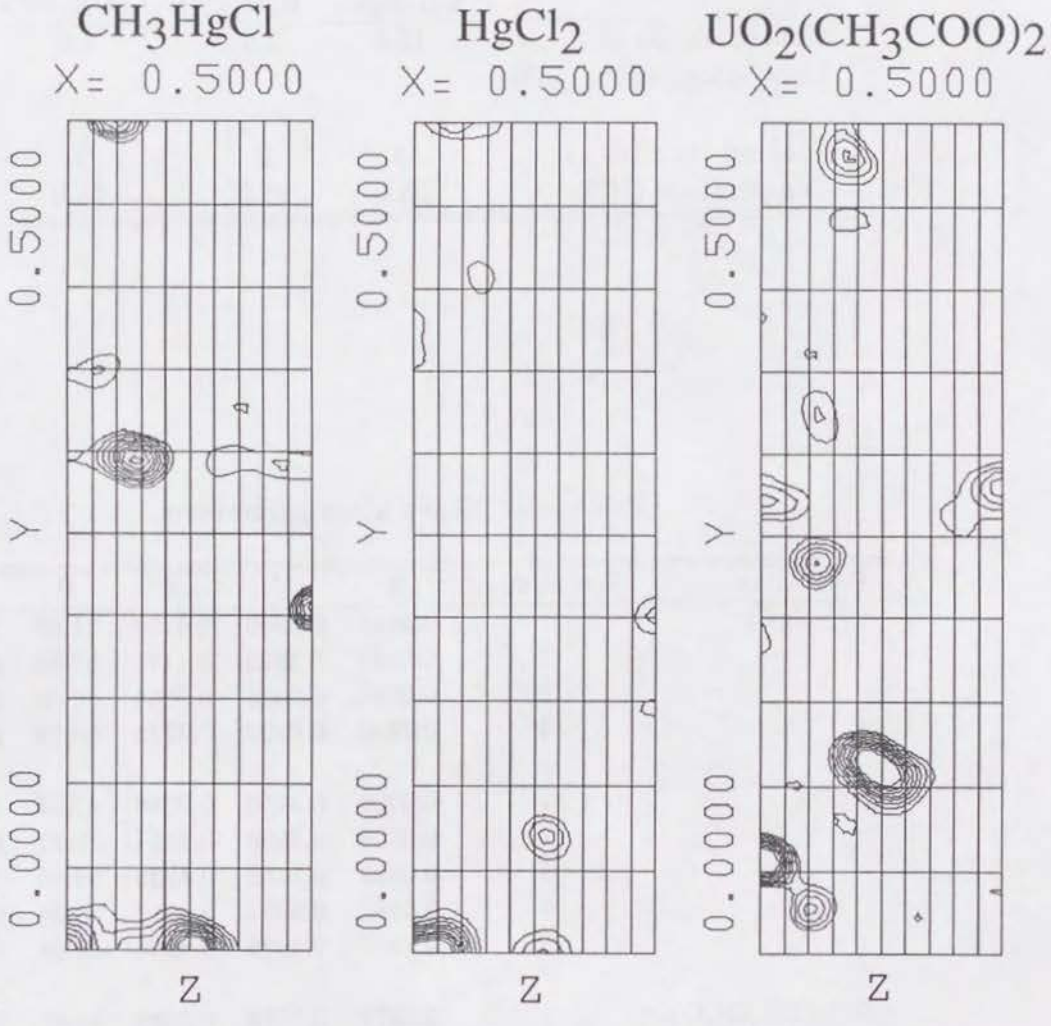


Figure 2.14: Difference Patterson maps on a Harker section (x=0.5).

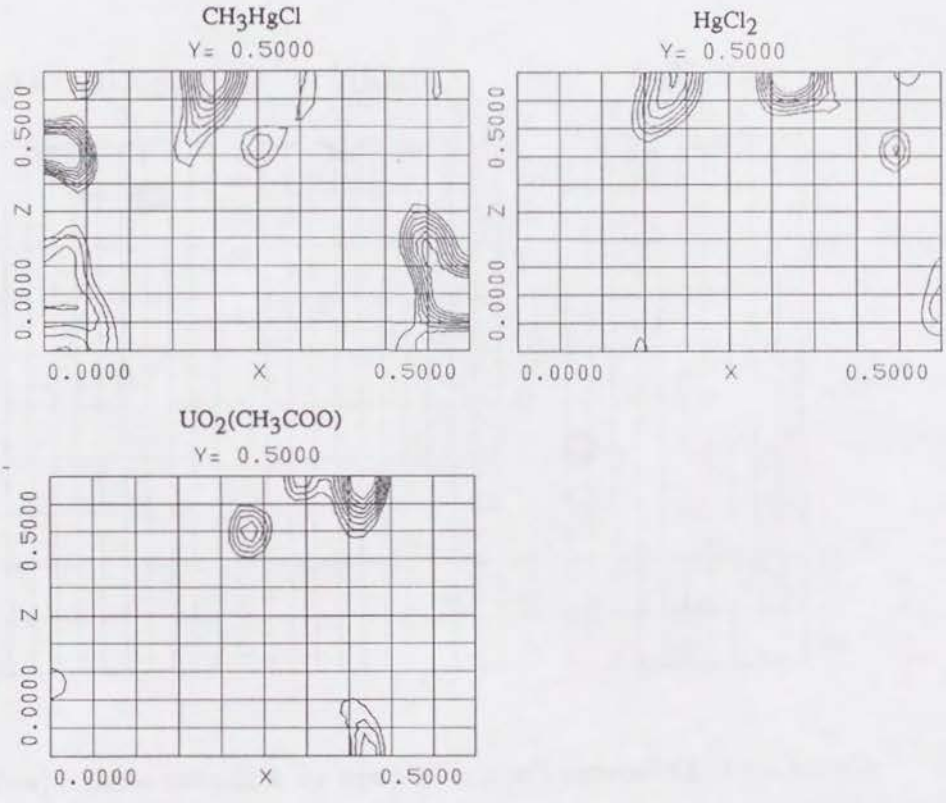


Figure 2.14: Difference Patterson maps on a Harker section (y=0.5).

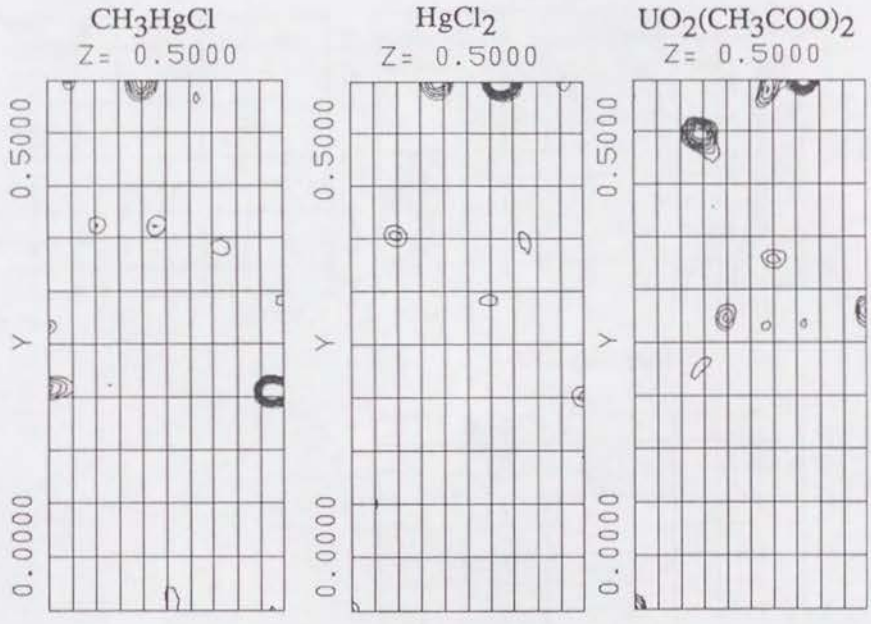


Figure 2.14: Difference Patterson maps on a Harker section (z=0.5).

solvent flattening at 2.8 Å resolution was made prior to the interpretation of the electron density map on the IRIS workstation. It is possible to mark directly on the mini-map with colored markers and to determine more easily the boundaries of molecules in the map than on the map in the graphic workstation. Z-sections of the map were plotted out at 2 Å intervals and copied on transparent OHP sheets (A4 size, Sumitomo 3M Inc. Ltd., Japan), and each of them was set on mounting flames (302×265 mm, Sumitomo 3M Inc. Ltd., Japan). The mini-map was then lighted up on a light box (Fujicolor light box 5000) and the probable molecular boundary, the N or C-termini, the molecular size, active site cleft and so on, were determined on the map. The probable contour of the main chain of the molecule was followed by marking the sections with colored markers.

The mini-map, which shows the electron density over the whole unit cell was calculated at 2.8 Å resolution using phased reflections, as shown in Figure 2.15. The map clearly shows molecular boundaries depicted in dashed lines. There were four molecules[9] of the alkaline protease in the unit cell, which was consistent with the previous estimation derived from the value of V_M. The molecule had an elongated elipsoidal shape, and seemed to consist of two structurally different domains, which were globular and elongated, respectively. The approximate size of the molecule was turned out to be 50×90×45 Å in the directions of *a*, *b* and *c* axes, respectively. However, much noise arose especially in the solvent region, and there remained some ambiguities mainly in loop regions on the surface of the molecule. Therefore, it was difficult to rationally interpret the electron density map.

Calculation of electron density maps

All procedures for calculating the electron density maps were performed using the program package PHASES [69]. The initial native electron density map was calculated at 2.8 Å resolution by a multiple isomorphous replacement (MIR) method. To improve the quality of the electron density map, a solvent-flattening technique [8] was applied to the map. After twenty-eight cycles, the interpretability of the map was improved. The solvent volume of the crystal was assumed to be 0.55. The procedures were as follows;

1. Computing the first solvent mask.
2. Computing four cycles of solvent flattening (using the first mask).
3. Computing the second solvent mask.
4. Computing four cycles of solvent flattening (using the second mask).

[9]Only two molecules were seen in the mini-map.

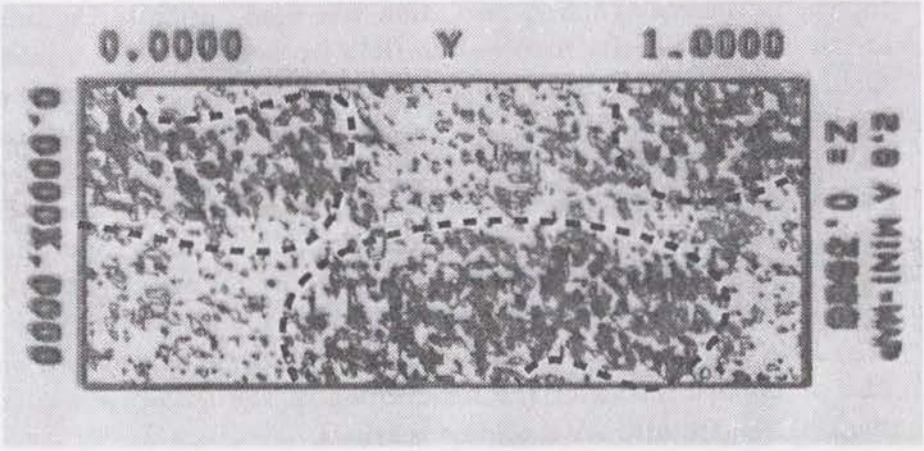


Figure 2.15: The mini-map of the alkaline protease at 2.8 Å resolution. This map was sectioned in the c axis. The map was calculated using the initial phased MIR-data (not solvent-flattened). The molecules can be seen in the regions enclosed by dashed lines. The approximate size of the molecule is $50\times90\times45$ Å. There is much noise in the solvent region.

5. Computing the third solvent mask.

6. Computing eight cycles of solvent flattening (using the third mask).

As a result, 16 cycles of solvent flattening and 3 cycles of mask calculation were performed iteratively.[10]

The averaged figure of merit increased from 0.571 to 0.759 during this procedure. The statistics for the map calculation are shown in Table 2.11. Examples of the electron density maps are shown in Figure 2.16 (a) the initial MIR map at 2.8 Å resolution, (b) the solvent-flattened map at 2.8 Å resolution, and (c) the final map refined at 2.0 Å resolution. The solvent flattening technique was generally not so effective for the main chain connectivity, but it greatly reduced the noise level in the solvent region, leading to the easier interpretation of the map.

Modeling on graphic workstation

The 2.8 Å electron density map, which had been solvent-flattened, was interpreted on the graphic workstation IRIS Indigo Elan using the program TURBO-FRODO [74, 75]. The initial model was built by superimposing each amino acid residue on the electron density map. To interpret am-

[10]doall.sh in the PHASES was used.

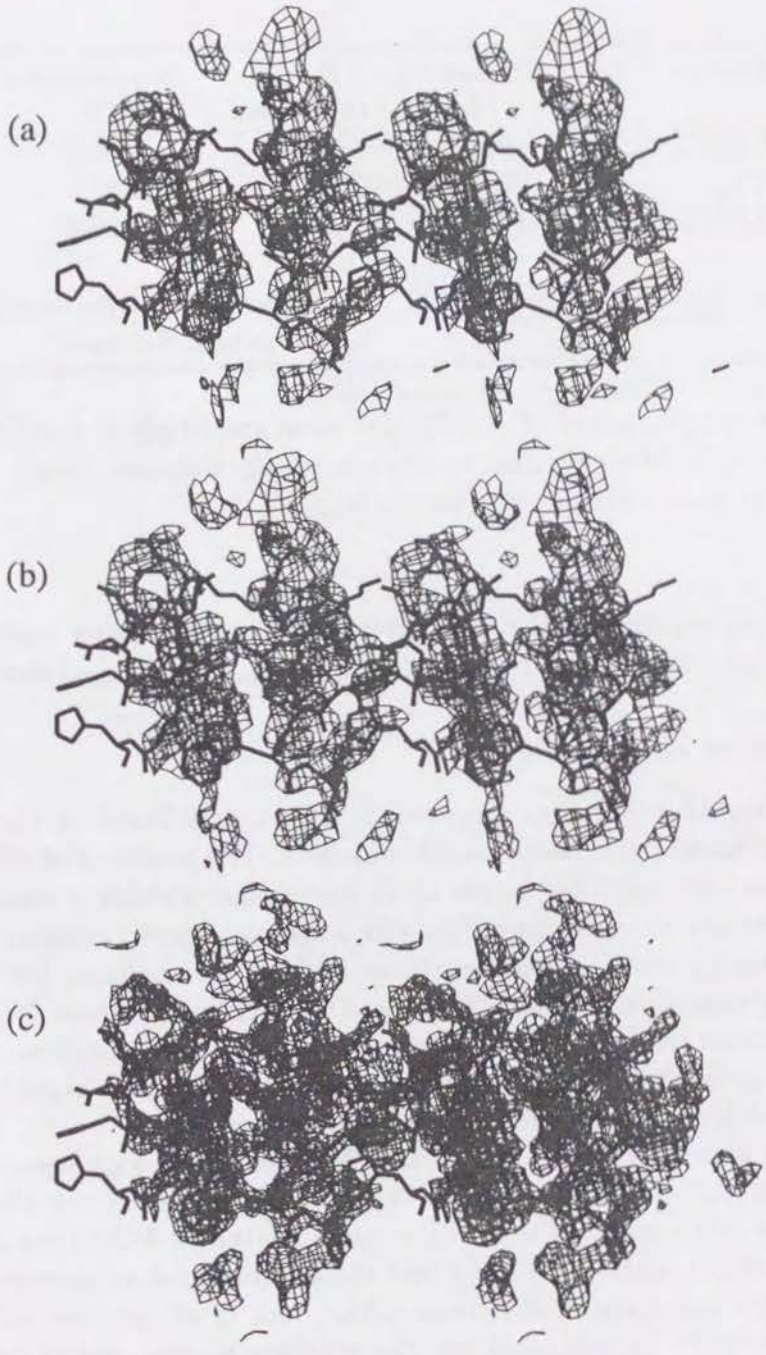


Figure 2.16: Examples of electron density maps around Q95–W103. (a) The initial MIR map drawn at 2.8 Å resolution, 1.5 $\sigma > I(\sigma)$, (b) after 28 cycles of solvent flattening at 2.8 Å resolution, 1.5 $\sigma > I(\sigma)$, (c) final electron density map at 2.0 Å resolution, 2.2 $\sigma > I(\sigma)$.

Table 2.11: Statistics for isomorphous replacement.

Derivative	Resolution (Å)	No. of reflections	completeness (%)	$R_{iso}^{a)}$ (%)	Phasig power[b)]
CH_3HgCl	10.0–2.8	15,662	89.6	24.1	1.83
$HgCl_2$	10.0–2.8	15,593	89.6	15.8	1.79
$UO_2(CH_3COO)_2$	10.0–2.8	15759	90.2	28.2	1.86
Averaged figure of merit[c)]		0.571 (not solvent flattened) 0.760 (solvent flattened)			

[a)] $R_{iso} = 100|I_{native}| - |I_{derivative}|/|I_{native}|$

[b)] Phasing power = $\langle F_H \rangle / \langle E \rangle$; the mean amplitude of the heavy-atom structure factor divided by the r.m.s lack-of-closure error.

[c)] Figure of merit = $|F_{hkl}(best)|/|F_{hkl}|$.

biguous regions in the map, difference Fourier maps with coefficients of $(2|F_{obs}| - |F_{calc}|)\exp(i\alpha_{calc})$ or $(|F_{obs}| - |F_{calc}|)\exp(i\alpha_{calc})$ were calculated.

Calcium binding sites

Calcium ions were incorporated into the model based on the height of electron densities and coordination numbers. The positions of calcium ions were further confirmed by means of an anomalous scattering measurement. It is well known that calcium ions, which are not directly coupled with enzymatic activity, bind to some enzymes such as thermolysin, [76], *Pseudomonas aeruginosa* elastase, and *Bacillus cereus* neutral protease [77]. The ions are associated with increase of thermal stability of these enzymes, although it has been pointed out that enhanced thermostability of calcium-bound enzymes cannot be attributed to a single determinant [43].

In general, however, removal of the calcium ions causes autolysis or irreversible loss of catalytic activity [49]. In the case of the alkaline protease, it was also suggested that the enzyme contained 1–2 atoms of calcium ions per mole of enzyme [78], and that the calcium binding prevented the enzyme from its autolysis. Furthermore, Morihara *et al.* pointed out that calcium ions must be incorporated into the synthesized calcium-free form of the alkaline protease secreted into a medium by *Pseudomonas aeruginosa*, and that the hypothetical calcium-free form might be more adapted for secretion than the calcium-bound one because of flexibility [79].

In the later study on protease B and protease C, which are homologous with the alkaline protease in their primary structure, it was shown that

they are synthesized as inactive higher molecular weight precursors with N-terminal extensions which are secreted into the external medium, where the divalent-cation-mediated activation occurs [80]. In this way, calcium ions play significant roles in enhancing thermal or structural stabilities and inducing activation of zymogens as as in the case of trypsin [81]. To determine calcium binding sites in the alkaline protease, native anomalous scattering effects derived from the bound calcium ions in the crystal were measured. To calculate each $\Delta|F|_{\text{anom}}$, which is the difference between the amplitudes of the structure factor for the reflections (h k l) and ($\overline{\text{h}}$ $\overline{\text{k}}$ $\overline{\text{l}}$) (Bijvoet pairs), the diffraction intensities of each Bijvoet pair were measured at 2.8 Å resolution by R-AXIS IIC (Rigaku Denki Co., Ltd., Tokyo, Japan). The measurement conditions are summarized in Table 2.12.

Table 2.12: Measurement conditions for anomalous scattering effects. One native crystal was used.

Crystal size	0.5×0.3×0.2 mm
Wavelength (Å)	1.5418 (Cu-$K\alpha$) Å
Measurement apparatus	R-AXIS IIC
Generator	RU-300
Focus	0.3×3.0, monochromated fine focus
X-ray power	40 kV, 100 mA
Rotation axis	b^*
Oscillation angle	3.0 deg.
Number of flames	33
Crystal to detector distance	160.385 mm
Detector two-theta angle	0.0114 deg.
Resolution	3.0 Å
Measurement temperature	20.0 °C

Refinement of the model

The initial model of the enzyme was refined by means of simulated annealing techniques using the molecular dynamics program, X-PLOR version 3.1 [82, 83, 84], installed on a CRAY Y-MP2E/264 supercomputer. In the preparation for simulated annealing refinement, several cycles of energy minimization were carried out to relieve strains or bad contacts in the initial model.

The slow-cooling protocol for the simulated annealing was then executed by reducing the temperature from 3,000 K to 300 K at intervals of by 25. At each temperature, 50 cycles of the refinement were carried out. During a whole series of cycles of the refinement, the W_A converged to about 300,000. After several cycles of the refinement, a manual rebuilding of the model was carried out using omitted difference Fourier maps. In the earlier stage of the refinement, classical protein structure parameters were used (called as 'param19x.pro' and 'toph19x.pro' in the X-PLOR package), and in the later stage, those of R. A. Engh and R. Huber were employed (termed 'parhcsdx.pro' and 'tohphscdx.pro') [85]. The resolution was gradually increased to 2.0 Å during iteration of the refinement and rebuilding. Water molecules were assigned to electron densities which appeared in the same positions both in $|F_{obs}| - |F_{calc}|$ ($> 3\sigma$) omit maps and $2|F_{obs}| - |F_{calc}|$ ($> \sigma$) omit-maps by manual interpretation. In the later stage of refinement, the positions of water molecules were automatically picked up using the program WATPEAK. The positions of the water molecules were confirmed by $2|F_{obs}| - |F_{calc}|$ and $|F_{obs}| - |F_{calc}|$ omit-maps. Calcium ions were incorporated into the model based on the height of electron densities and the coordination number. The positions of calcium ions were further confirmed by means of an anomalous scattering measurement. Individual temperature factors were also refined at 2.5 Å resolution. At 2.3 Å resolution, the b^*-rotation native data set was merged with the c^*-rotation set. The whole course of the refinement strategy is shown in Figure 2.17.

The refinement procedure was not straightforward because crystallographic R-values converged to slightly above 20.0 % at 2.0 Å resolution on the later cycles of the refinement as shown in Table 2.17, and the r.m.s.deviations[11] from the ideal values of the bond lengths and residual angles of the model have converged to 0.017 Å and 3.122°, respectively. These values were not so good for the model refined at 2.0 Å resolution. This is because the protein structure parameters used were exchanged to those of Engh and Huber, which can be expected to offer a better model in its geometries, and because observed the structure factor amplitudes were scaled to those from the refined model. After the exchange and the scaling, the R-value reduced to 19.3 % at 2.0 Å resolution and the r.m.s. deviations for bond lengths and angles were improved to 0.001 Å and 1.595°, respectively. In the final cycle of the refinement, the free R value (R_{free}) [86] was calculated. The final coordinates of the alkaline protease has been processed by the Brookheaven Protein Data Bank with the ID code **1AKL** and will be available soon by the anonymous FTP. The strategy and the course of refinement is summarized in Figure 2.18.

To estimate the deviation from a Luzzati plot [87] and to determine the

[11] root-mean square deviation.

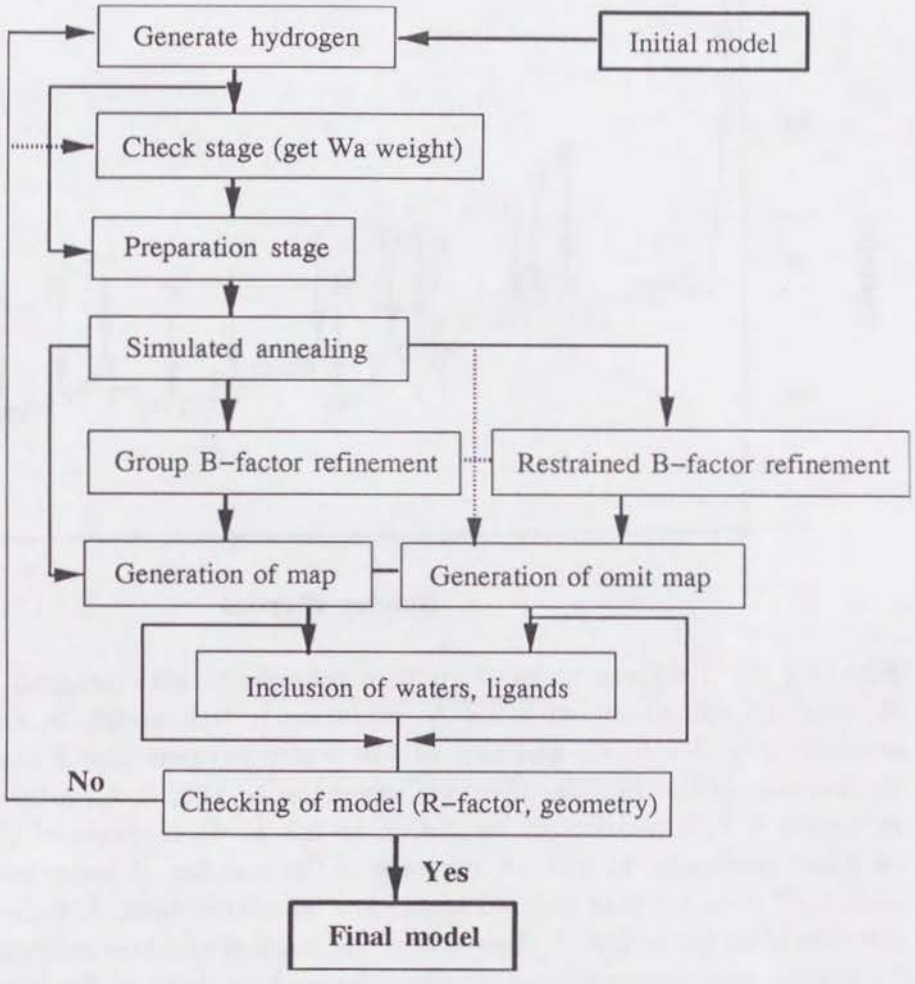


Figure 2.17: Refinement strategy of the model of the alkaline protease.

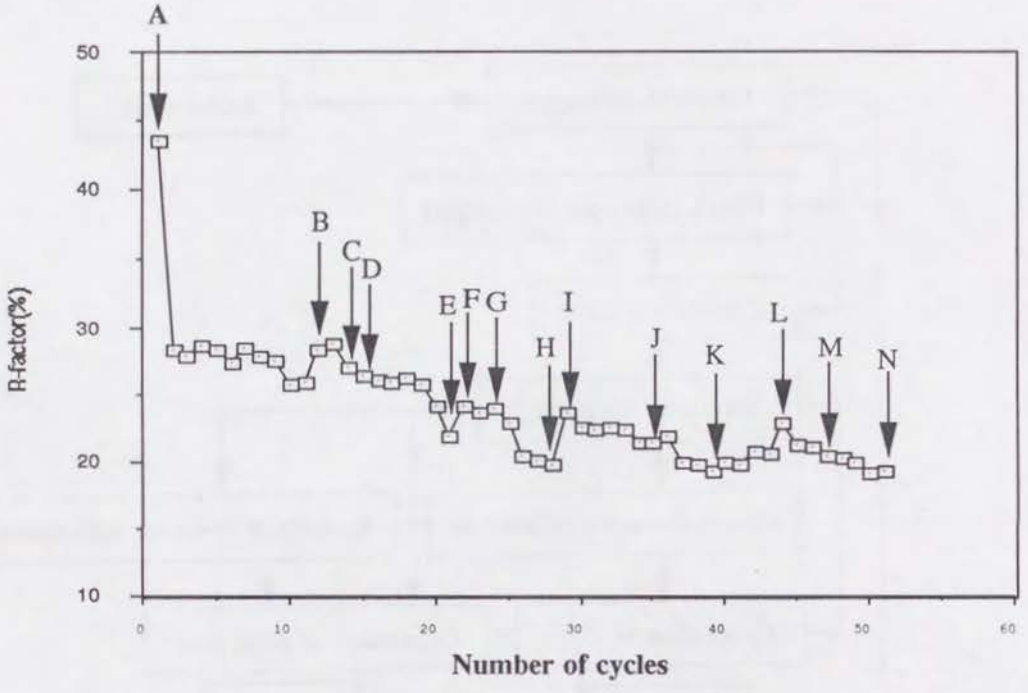


Figure 2.18: Progress in the structure refinement with program X-PLOR: A, start of refinement with 2.8 Å resolution initial model; B, extension of resolution to 2.5 Å; C, addition of 110 water oxygens and 8 calcium ions; D, increase of the number of water molecules to 183; E; temperature factor refinement; F, extension of resolution to 2.3 Å; G, increase of the number of water molecules to 253; H, increase of the number of water molecules to 303; I, b^* data set was merged to the initial native data; J, decrease in the number of water to 248; K, increase of the number of water molecules to 301; L, protein structure parameters were changed to those of Engh and Huber; M, the native data set was scaled to the model; N, increase in the number of water molecules to 331.

absolute scale K and an overall temperature factor B, the Luzzati plot and Wilson plot [88] for the alkaline protease were calculated. They are shown in Figure 2.19 and Figure 2.20, respectively. The deviation from a Luzzati plot was estimated by the plot. The absolute scale factor K and tempera-

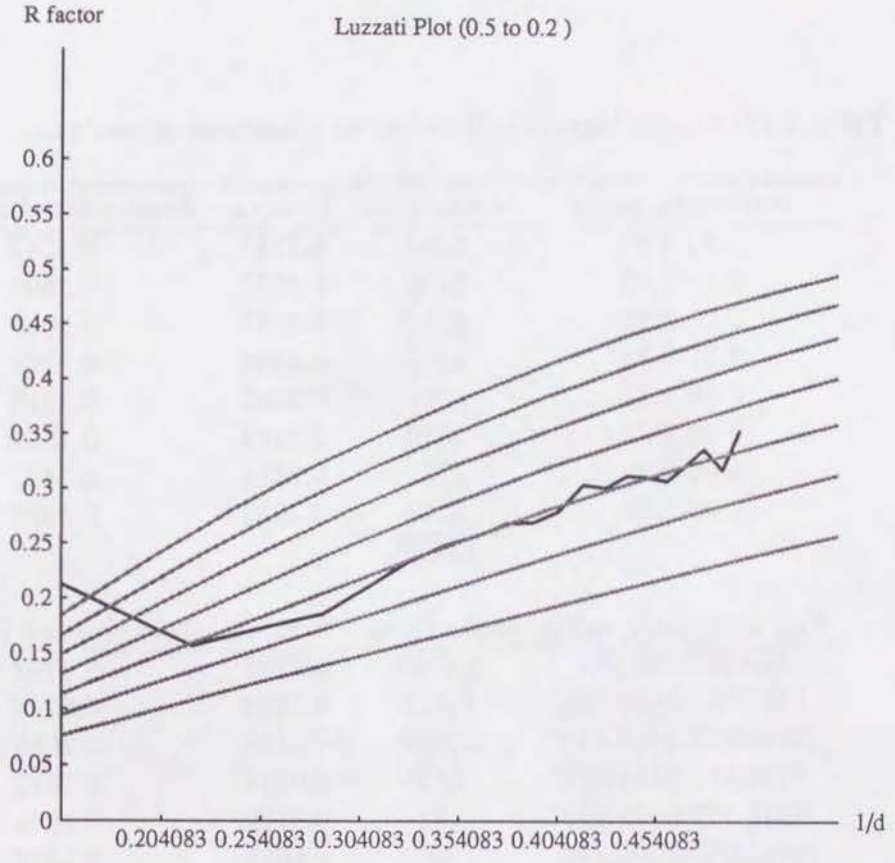


Figure 2.19: Luzzati plot for the refined model of the alkaline protease at 2.0 Å resolution. The plot is drawn based on 33,698 reflections up to a resolution of 2.0 Å. The estimated deviation from the Luzzati plot is 0.3.

ture factor B of structure factor amplitudes $|F_{abs}|$ can be determined by the following equation,

$$|F_{abs}|^2 = K|F_{osb}|\exp(-2B\sin(\theta))/\lambda^2$$

The final model statistics and refinement parameters are summarized in Table 2.14. A Ramachandran plot, depicted by the program PROCHECK [89], is shown in Figure 2.21. All residues are observed in the allowed region of the Ramachandran plot, and 87.6 % of the total residues (353 residues) lie

Table 2.13: Crystallographic R-values as a function of resolution or intensity.

resolution range	reflections	R-value	Accumulated R-value
3.86-8.00	6041	0.1547	0.1547
3.13-3.86	5636	0.1671	0.1597
2.75-3.13	5110	0.2147	0.1695
2.51-2.75	4591	0.2308	0.1762
2.33-2.51	4011	0.2495	0.1816
2.20-2.33	3310	0.2689	0.1861
2.09-2.20	2751	0.2814	0.1897
2.00-2.09	2248	0.3053	0.1928
	33,698		

F_{obs} amplitude range	reflections	R-value	Accumulated R-value
25.843-1353.265	23,943	0.2798	0.2798
1353.265-2680.688	6,512	0.1628	0.2297
2680.688-4008.110	2,359	0.1159	0.2058
4008.110-5335.532	669	0.0934	0.1972
5335.532-6662.955	151	0.0709	0.1944
6662.955-7990.377	47	0.0628	0.1933
7990.377-9317.800	11	0.0447	0.1930
9317.800-10645.222	6	0.0518	0.1928
	33,698		

Table 2.14: Final model statistics and refinement parameters at 2.0 Å resolution.

Data statistics	
number of reflections	33,698
resolution range (Å)	8.0–2.0
completeness for range	0.751
data cutoff	2.0 σ
Refinement statistics	
R-value[a)]	0.193
free R-value[b)]	0.225
free R-value test set	10.0 %
mean B-value (Å^2)	26.55
estimated deviation from Luzzati plot (Å)	0.3
r.m.s.d in	
bond length	0.010 Å
residual angles	1.595 deg.
dihedral angles	26.12 deg.
improper angles	1.201 deg.
Number of	
protein atoms	3,504
water molecules	331
zinc ion	1
calcium ions	8

[a)] $R - \text{value} = \Sigma||F_{obs}(h\ k\ l)| - |F_{calc}(h\ k\ l)||/\Sigma|F_{obs}|$.

[b)] $\text{free R value} = \Sigma_{(h\ k\ l)}||F_{obs}(h\ k\ l)| - |F_{calc}(h\ k\ l)||/\Sigma_{(h\ k\ l)}\int|F_{obs}|$.

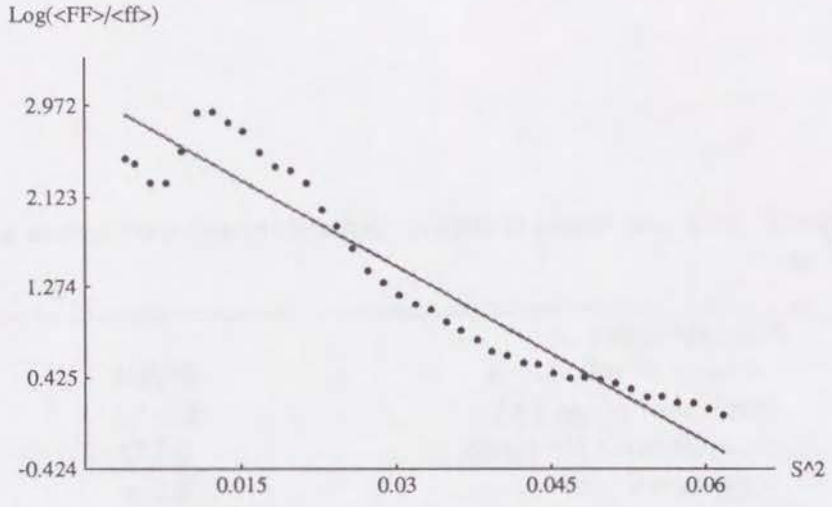


Figure 2.20: A Wilson plot based on 13,153 reflections up to the level of 2.0 Å resolution. The value of the temperature factor B and scale factor K were determined to be 27.78 $Å^2$ and 0.0433 from the best fitted line,respectively.

in the most favored region, 11.7 % (47 residues) are present in the additional allowed region and 0.7 % (3 residues) are in the generally allowed regions. No non-glycine residues fall into the disallowed regions. The average temperature factors of the main and side chains are plotted against the residue numbers as shown in Figure 2.22. The absolute deviation from mean χ_1, χ_2, ω torsions , and ζ torsions, and secondary structures assignment by the method of W. Kabsch and C. Sander [90] are shown in Figure 2.23.

2.3 The N-terminal catalytic domain

A ribbon drawing of the alkaline protease and a folding diagram are depicted in Figure 2.24 and Figure 2.25, respectively. The enzyme has an elongated ellipsoidal shape with approximate dimensions of 90×42×35 Å. The enzyme has twelve α-helices, twenty five β-strands, one catalytically essential zinc ion and eight calcium ions. The structure comprises two structurally distinct domains: a globular proteolytic (or N-terminal) domain whose folding topology is very similar to those of other metzincins as discussed later, and a C-terminal domain where β-strands are dominantly present. Apart from the N-terminal region including a α-helix (α1 in Figure 2.25), which is associated with the C-terminal domain, the proteolytic domain comprising residues 18–250 the catalytically essential zinc ion in the active center. The domain has

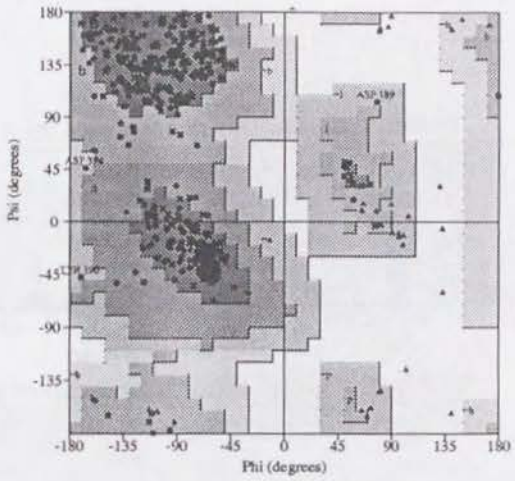


Figure 2.21: A Ramachandran plot of the alkaline protease refined at 2.0 Å resolution. Triangles indicate glycine residues and squares, non-glycine residues. Tyr190 is in a flexible loop region. The total number of residues is 470, 87.6 % of which lie in the most favored regions [A,B,L]. 11.7 % of the total residues are in the additional allowed regions [a,b,l,p], and 0.7 % fall into the generously allowed regions [~a,~b,~l,~p].

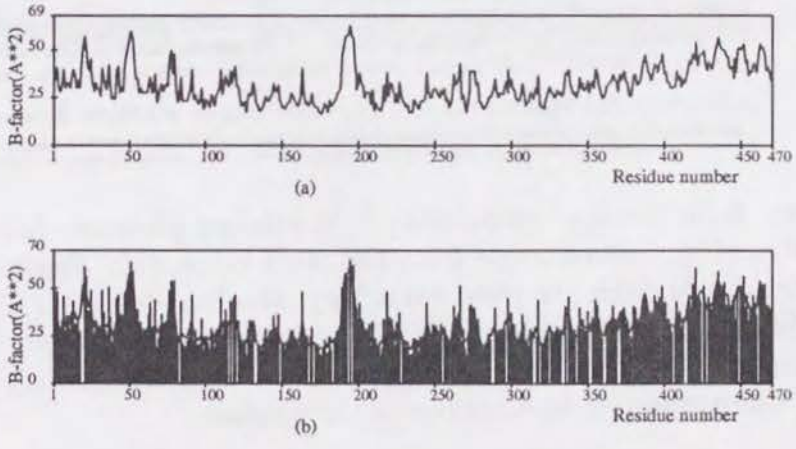


Figure 2.22: Temperature factor (B Å^2) plots against residue number. (a) that of the main chain and (b) that of the side chain.

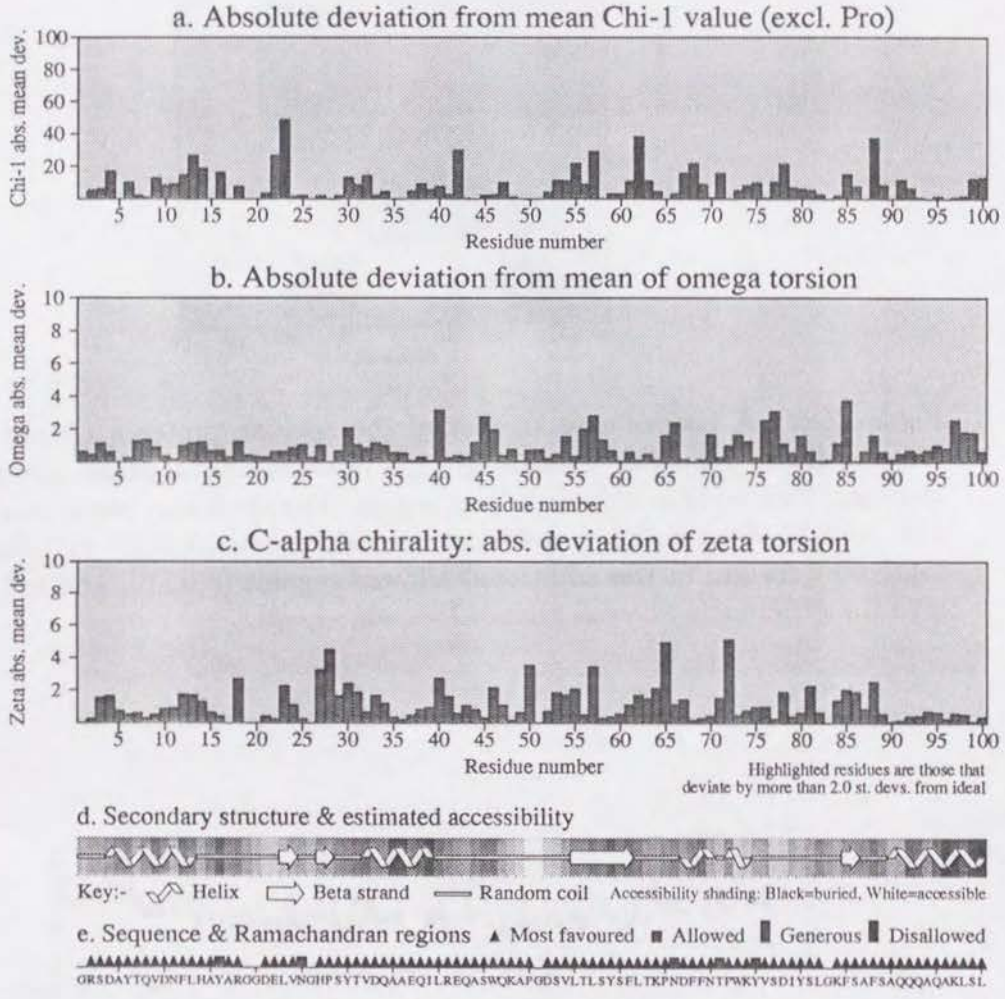


Figure 2.23: Residue properties of the alkaline protease. (a): Absolute deviation of χ_1 torsion angle from the ideal value; (b): absolute deviation ω torsion angle from the ideal value; (c): absolute deviation of ζ torsion angle (defined by the atoms C_α-N-C-C_β) from the ideal value; (d): schematic representation of secondary structure and estimated accessibility assignment; (e): Ramachandran plot regions of the residues.

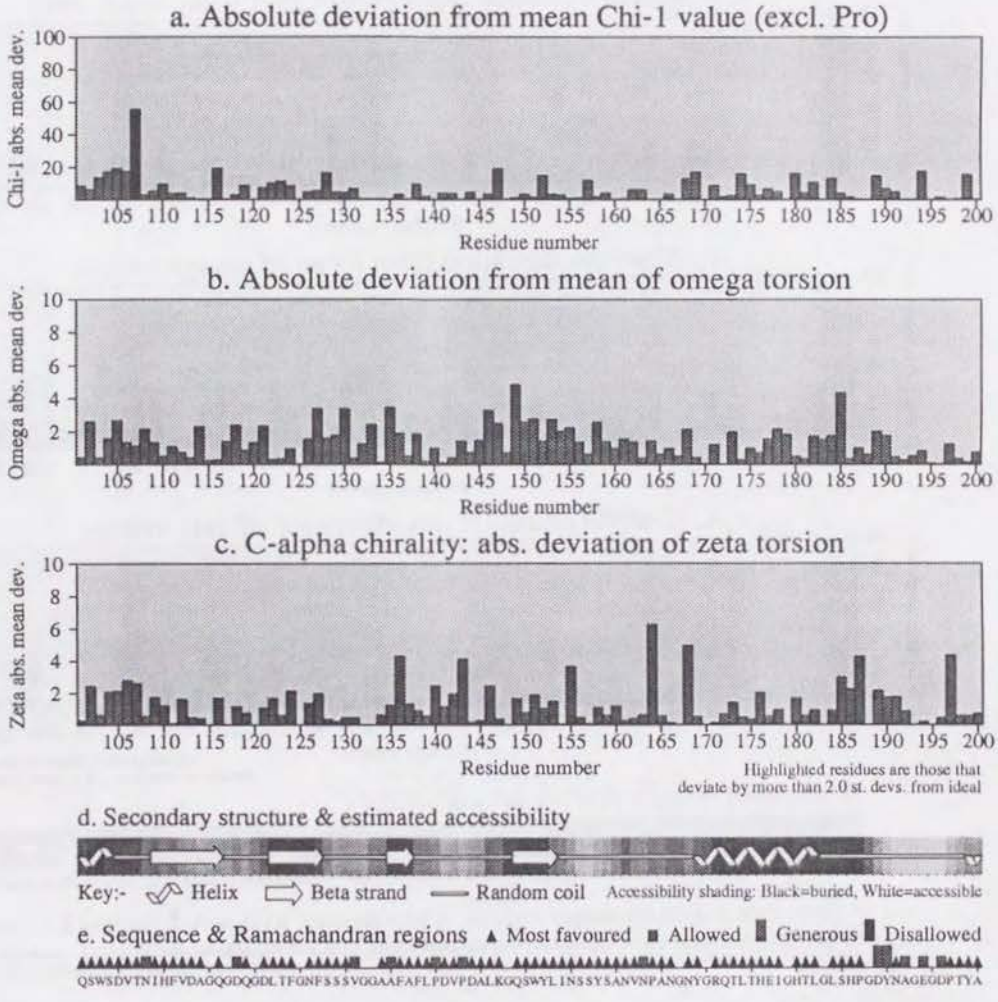


Figure 2.23: —*continued.*

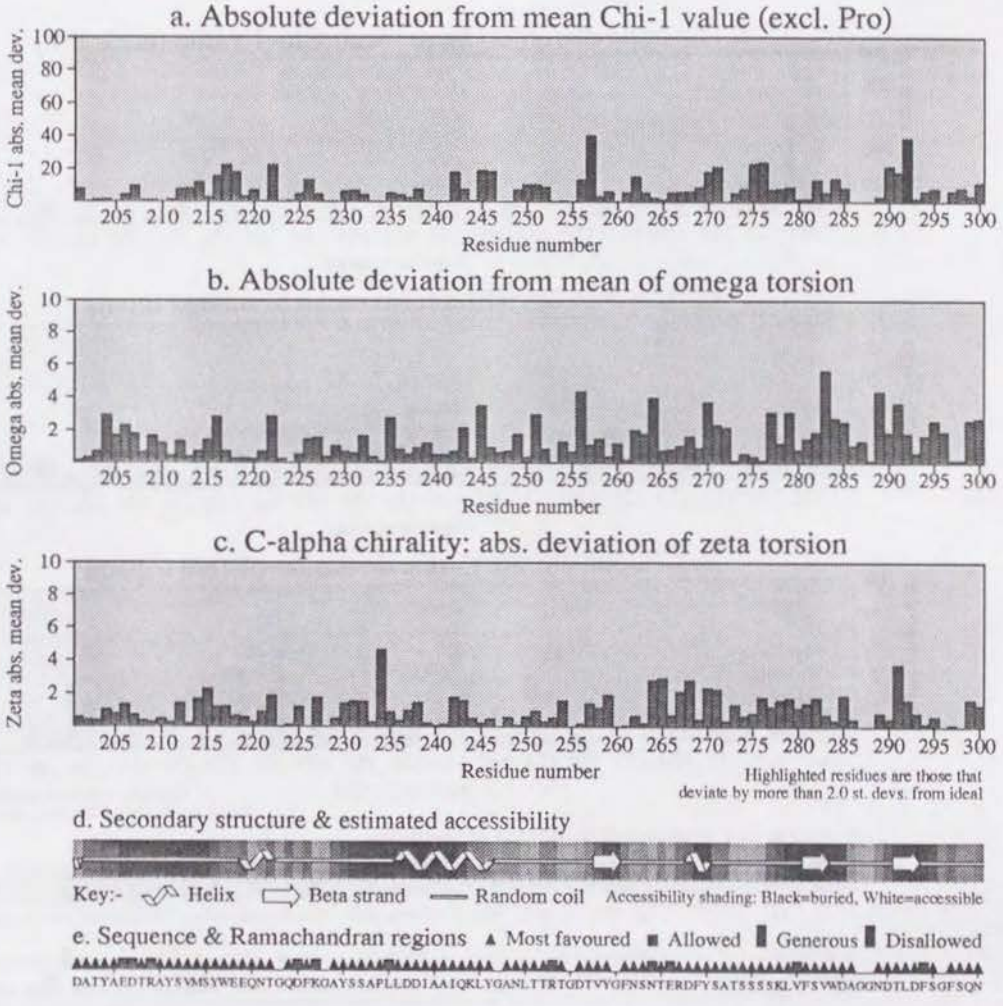


Figure 2.23: —*continued.*

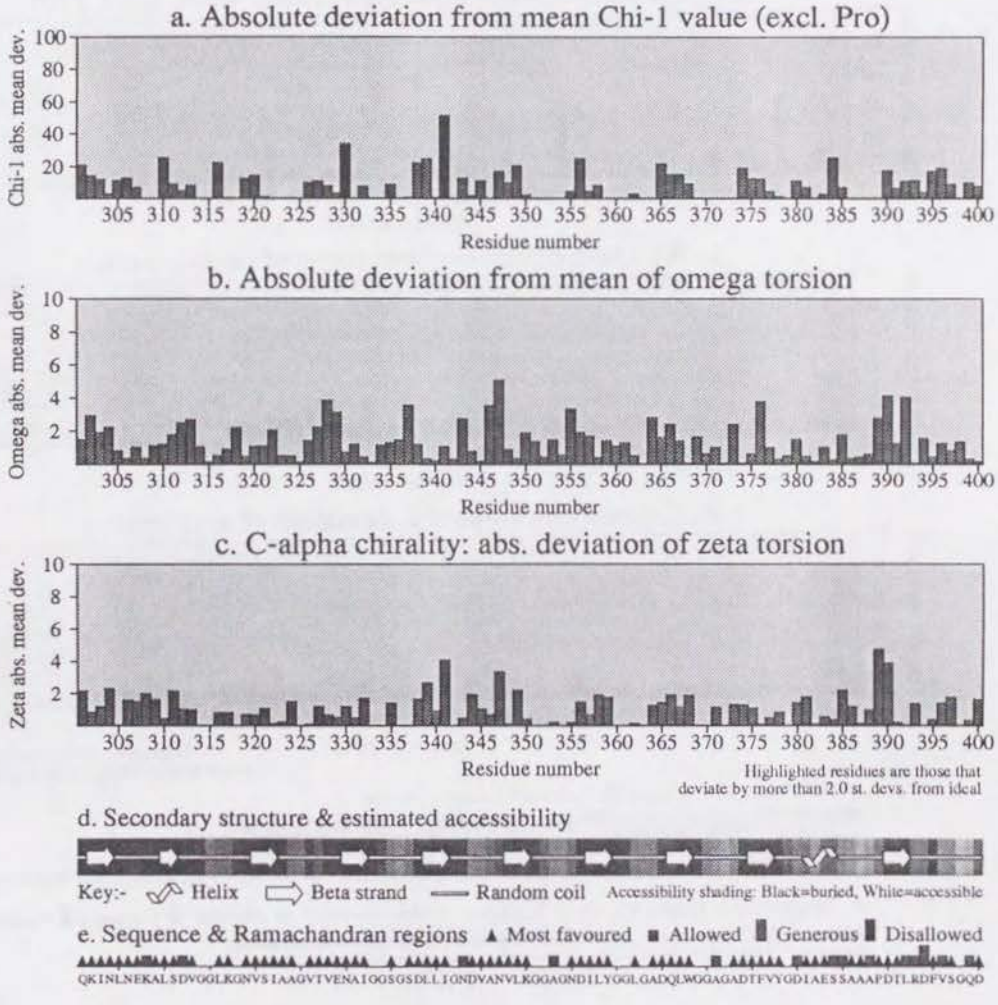


Figure 2.23: —*continued.*

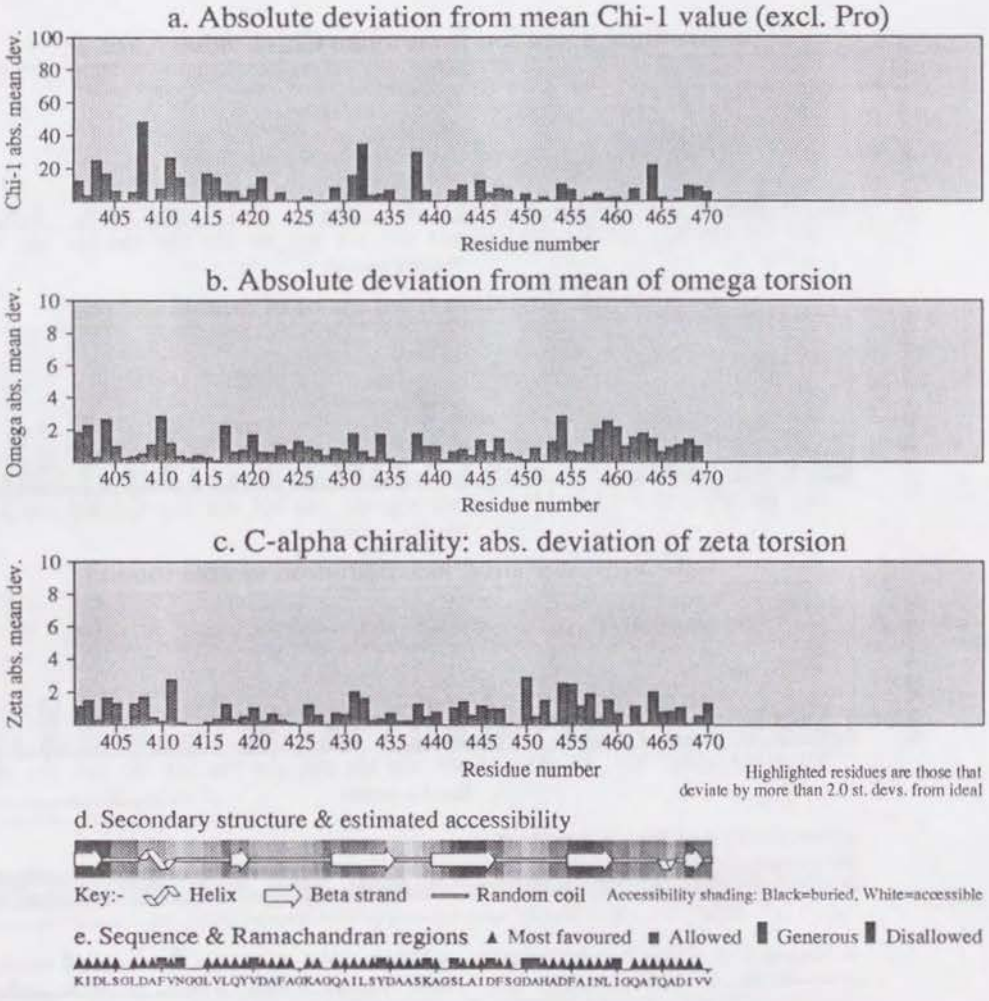


Figure 2.23: —*continued.*

a globular shape, with approximate dimensions of 35×38×38 Å, as depicted in Figure 2.26. It further consists of two subdomains: the "upper" and the "lower" subdomains. They are separated by the active site cleft, which horizontally runs through the whole width of the enzyme. The "active site helix" $\alpha 6$ comprising residues Y169–L182, which is inserted between $\beta 4$ and $\beta 5$, lies at the bottom of the cleft with the zinc ion.

The upper subdomain partly consists of five β-strands, $\beta 3$, $\beta 4$, $\beta 5$, $\beta 6$ and $\beta 7$, which are connected by coils in the manner, -1x, +2, +2, -1 with $\beta 6$ antiparallel to the others [92]. The β-strands make a twisted β-pleated sheet which serves as the upper wall of the subdomain. Behind the $\alpha 6$, an α-helix $\alpha 5$, consisting of residue A90–S104, runs obliquely from top to bottom. A coil consisting of Y75–K89, which connects $\alpha 4$ and $\alpha 5$, makes a bulge over the β pleated sheet being the upper lim of the active cleft.

Zinc binding site

The lower subdomain mainly consists of coils and α-helices, $\alpha 7$ to $\alpha 9$, making up the lower rim of the active site cleft. The long α-helix $\alpha 9$, consisting of residue L236–Y246, lies nearly horizontally from right to left, serving as a half of the bottom of the active site cleft. A long coil, consisting of residues G223–L235, makes the other half of the bottom connecting $\alpha 8$ and $\alpha 9$.

The $2|F_{obs}| - |F_{calc}|$ electron density map superimposed on the skeletal model around the zinc ion is drawn in Figure 2.28 (a). The map was calculated using the phase angles from the final refined model, contoured in 1.8 σ level. Small globular-like electron densities correspond to bound water molecules around the active site. The active site cleft has a highly hydrophilic environment and approximately 28 water molecules, as shown in Figure 2.27. This may contribute to an increase in the accessibility of substrates. Most of the bound water molecules are replaced with ligands which bind to the enzyme. This is consistent with the case of the complexes between thermolysin and the transition state analoques [93].

The geometries around the zinc ion are shown in Table 2.15. In the zinc ligation, the imidazole nitrogens of His176 and H186, and the water oxygen are located on the vertices of the triangular base plane which accommodates the active site zinc ion. The imidazole nitrogen of His180 and the phenolic oxygen of Tyr216 lie perpendicular to this triangle and occupy the vertices of the bipyramid. Tyrosine residues equivalent to Tyr216, which serves as the fifth zinc ligands in the alkaline protease, are conserved in the sequences of all members of the serralysins and the astacins, as shown in Figure 2.36, on page 75, but not in the other three superfamilies. The distance between the zinc ion and the phenolic oxygen of Tyr216 is somewhat long to make a coordinate bond with each other. Therefore, the side chain of Tyr216 is

Figure 2.24: A ribbon drawing of alkaline protease from *Pseudomonas aeruginosa* IFO3080, depicted with the program MOLSCRIPT [91]. In the figure, α-helices are denoted by helical ribbons colored in red, β-strands by blue arrows, the active site zinc ion by a green sphere, and calcium ions by magenta balls, respectively.

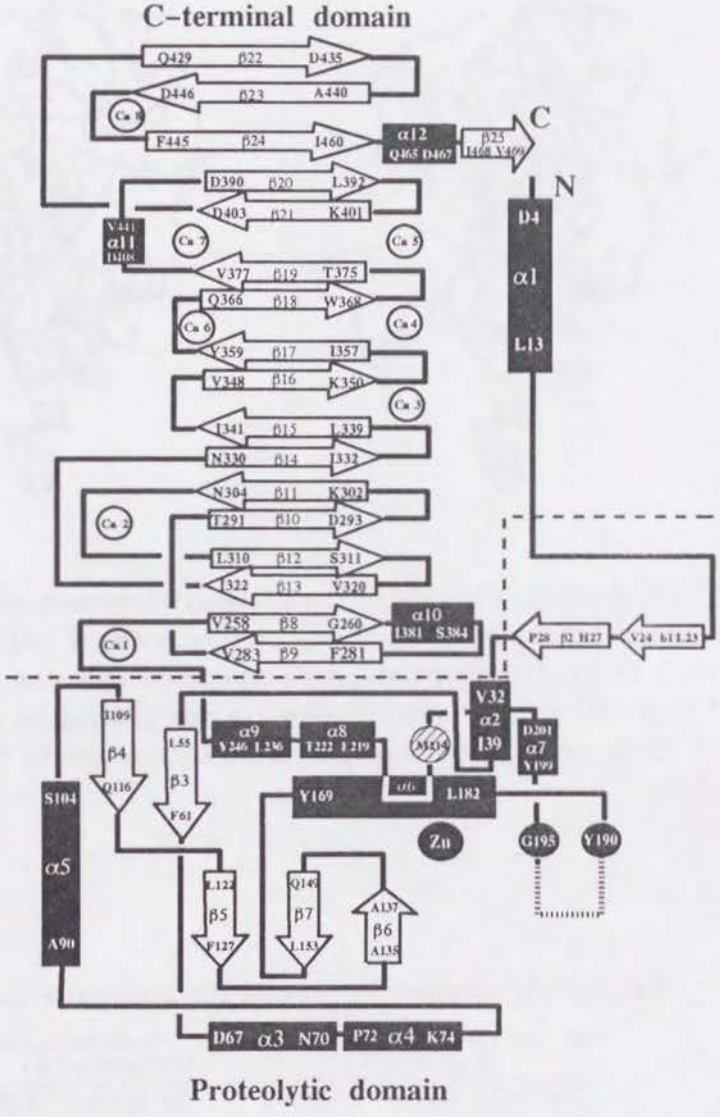


Figure 2.25: A schematic representation of the secondary structure elements of the alkaline protease. All of the secondary elements are depicted with corresponding one-character codes. α-Helices are depicted as filled black cylinders, β-strands as arrows. The active site zinc is drawn as a black ball with a code 'Zn', calcium ions as white balls. The conserved Met-turn is located around M214, drawn as a striped ball. A flexible loop region is presented as a doted line.

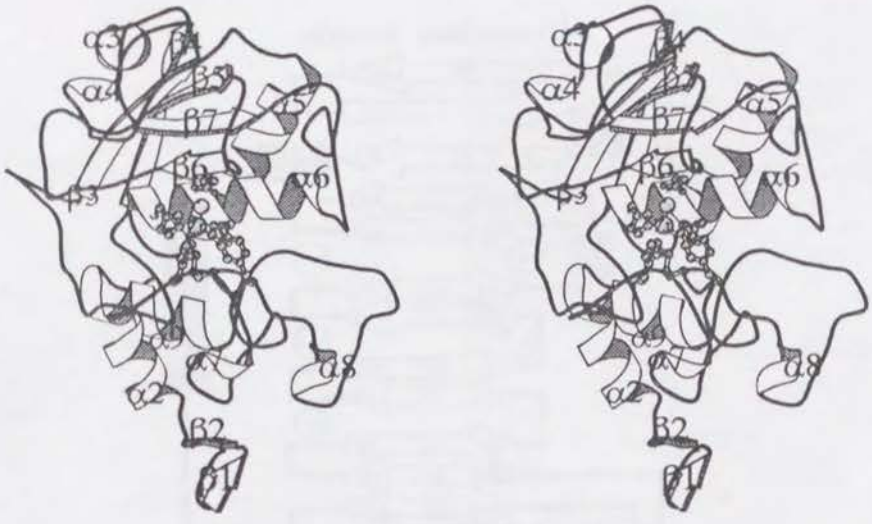


Figure 2.26: A stereo drawing of the proteolytic domain of the alkaline protease. The relative orientation of this picture is about 180 ° to the molecule in Figure 2.25 rotated about an axis perpendicular to the paper plane. Helical ribbons and arrows indicate α-helices and β-strands, respectively. The numbers on secondary structure elements correspond to those depicted in Figure 2.25.

Table 2.15: Geometries around the active site zinc ion.

Angles (°)		Distances (Å)	
Nϵ186-Zn-Nϵ180	95.03	Nϵ180-Zn	2.20
Nϵ186-Zn-H_2O	120.97	OH216-Zn	3.02
Nϵ176-Zn-H_2O	118.53	Nϵ176-Zn	2.36
Nϵ180-Zn-H_2O	114.34	Nϵ186-Zn	2.22
Nϵ180-Zn-Nϵ176	99.61	Nϵ186-Zn	2.16
OH216-Zn-H_2O	63.27		
OH216-Zn-Nϵ176	82.93	OH216-H_2O	2.86
OH216-Zn-Nϵ186	85.89	O$\epsilon_1$177-H_2O	2.88
OH216-Zn-Nϵ180	177.24	O$\epsilon_2$177-H_2O	3.53

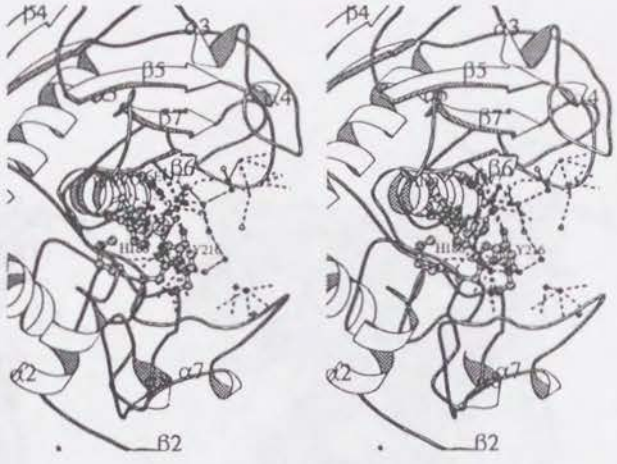


Figure 2.27: A stereo drawing of the active site with bound water molecules. Small balls represent water molecules with dashed lines which show hydrogen bonds. There are approximately 28 water molecules inside the active site cleft which make a hydrogen bond network.

able to rotate relatively freely around the zinc ion. As discussed later, the Tyr216 plays an important role in binding ligands, making hydrogen bonds with them.

A flexible loop region

The loop region of Y190–G195 (Y190-N191-A192-G193-E194-D195) in the native alkaline protease, which is depicted as a dashed line in Figure 2.25, exhibited an unclear electron density. The average temperature factor in the region diverges to more than 70 $Å^2$, as shown in Figure 2.22, page48 which implies that this region disordered. Interestingly, the sequence of the flexible loop region in the serralyins is highly conserved as shown in Figure 2.29, although it is common in proteins that flexible loops their molecular surfaces are less conserved than functionally important portions. This implies that the flexible loop region may play an important role in catalytic function.

2.3.1 Functions of the C-terminal domain of the alkaline protease

The catalytic (N-terminal) domains of metzincins have a basically common folding topology as discussed

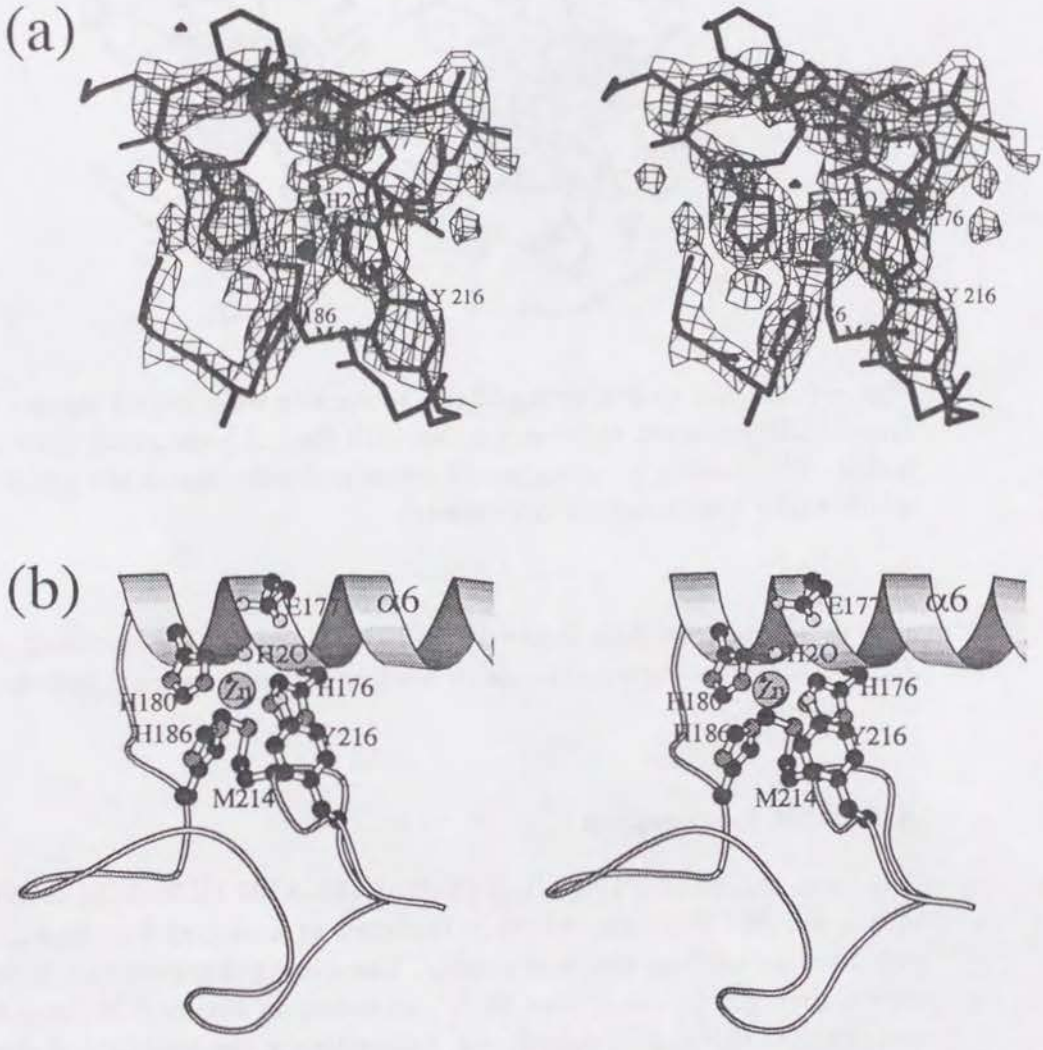


Figure 2.28: A stereo drawing of the area around an active site zinc ion. (a) The $2|F_{obs}| - |F_{calc}|$ map at 2.0 Å resolution, superimposed on the skeletal model. The contours are drawn at 1.8 σlevel. The zinc ion has five ligands, His176, His180, H186, Tyr216 and one H_2O molecule. (b) A ball-and-stick representation around the active site. The conserved Met-turn including Met214 is located beneath the active zinc ion.

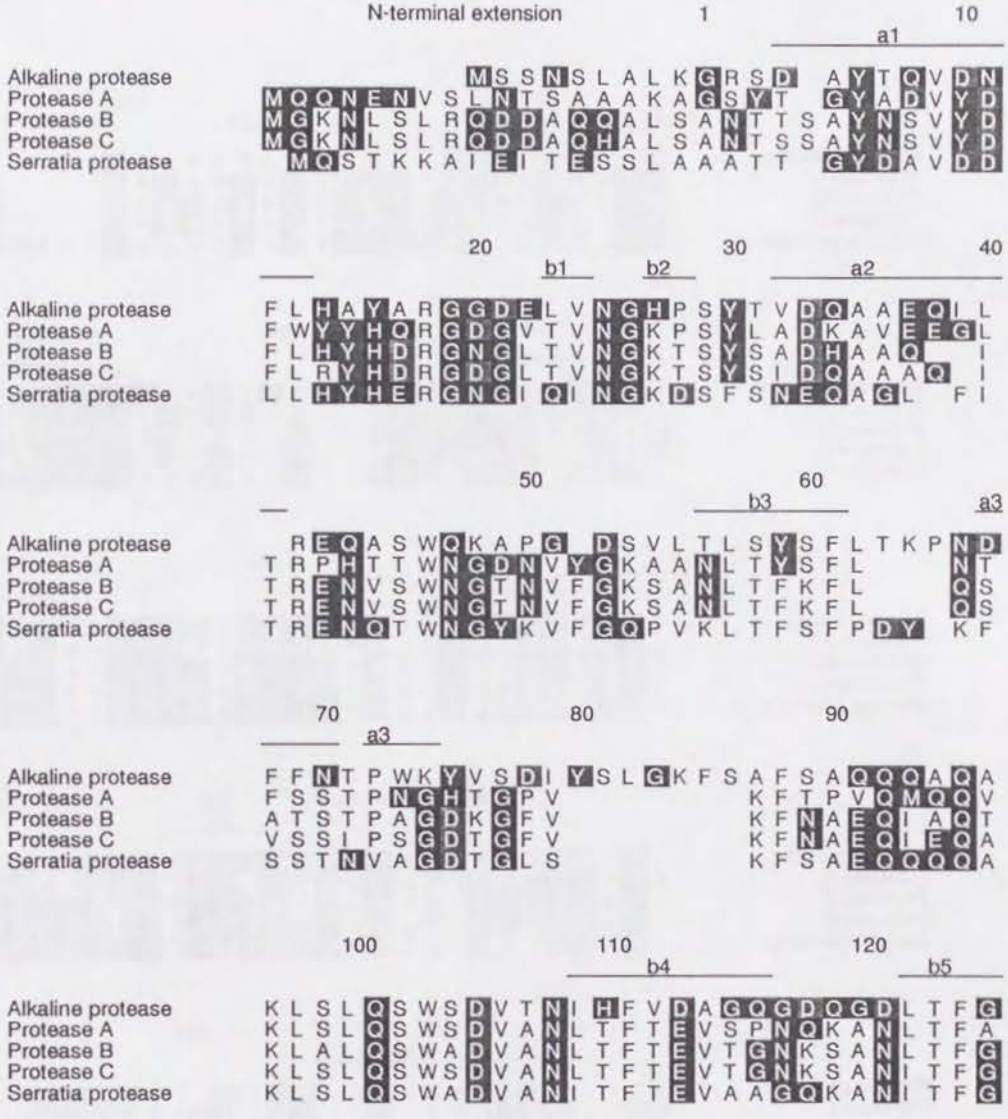


Figure 2.29: Sequence alignment of serralysins. The sequence numbers correspond to the alkaline protease. Zinc ligands are indicated by '*'s. The conserved zinc binding motif around His180, a flexible loop region, a conserved Met-turn, and repeated GGXGXD motifs are boxed or indicated. The alignment was performed with program package AMPS [94], and depicted using program ALSCRIPT [95].

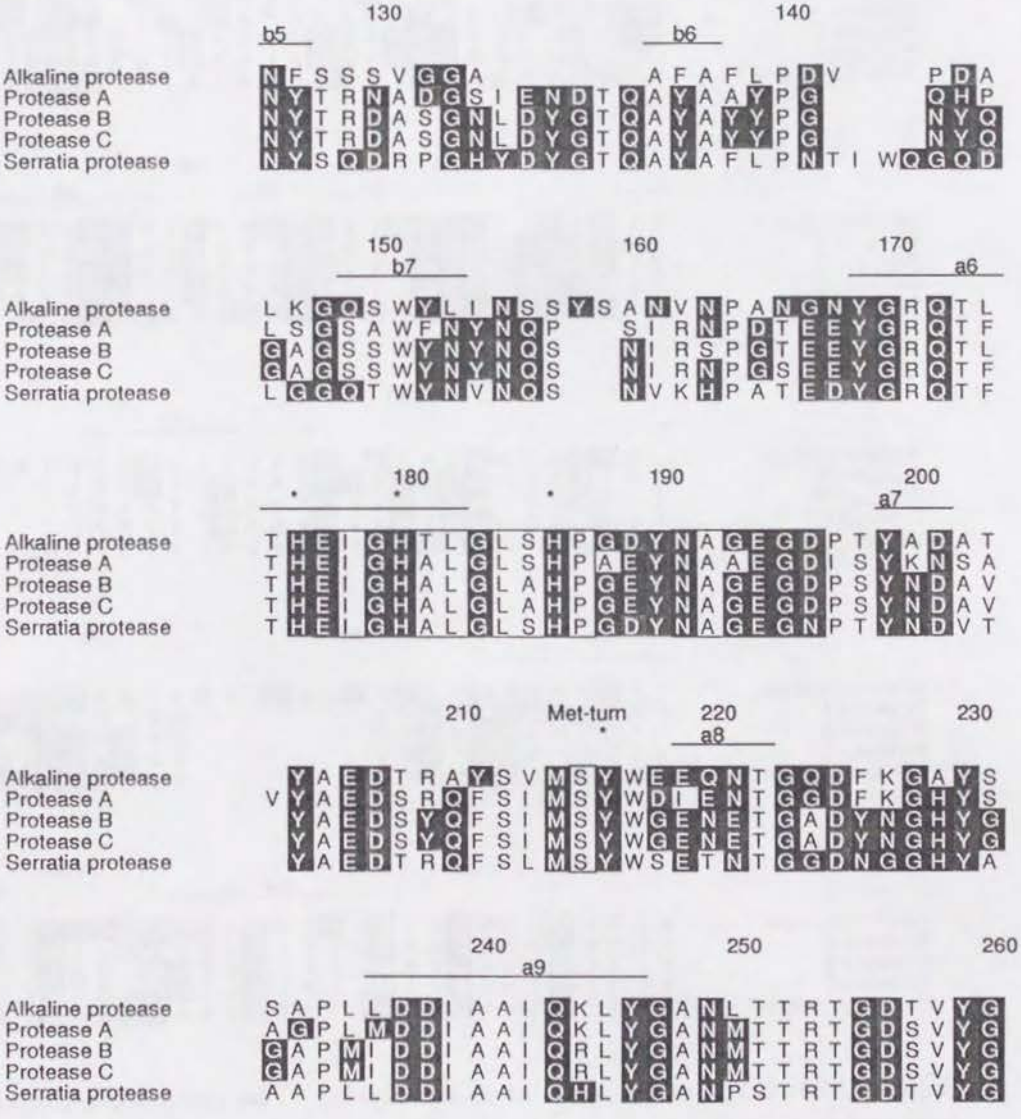


Figure 2.29: —*continued.*

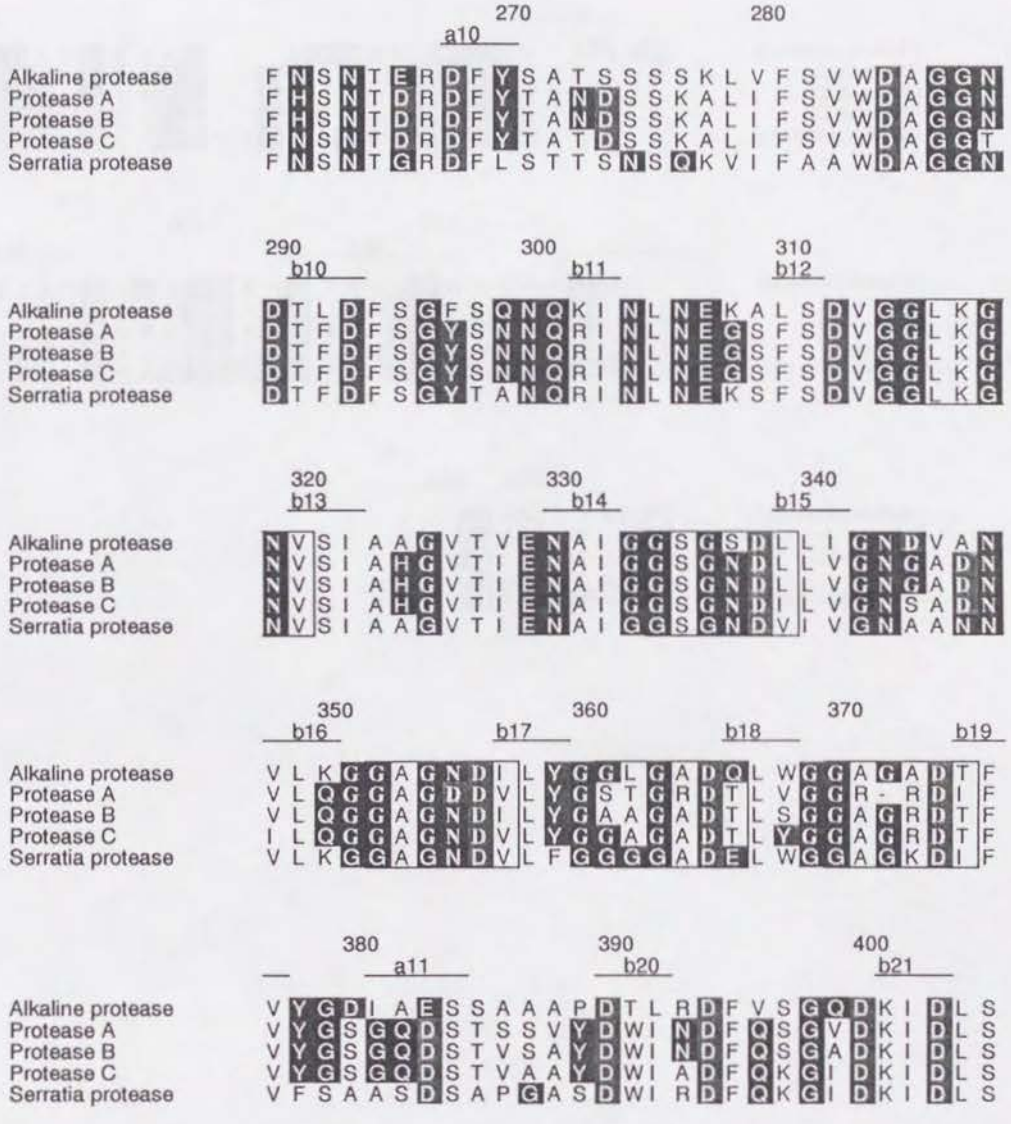


Figure 2.29: —*continued.*

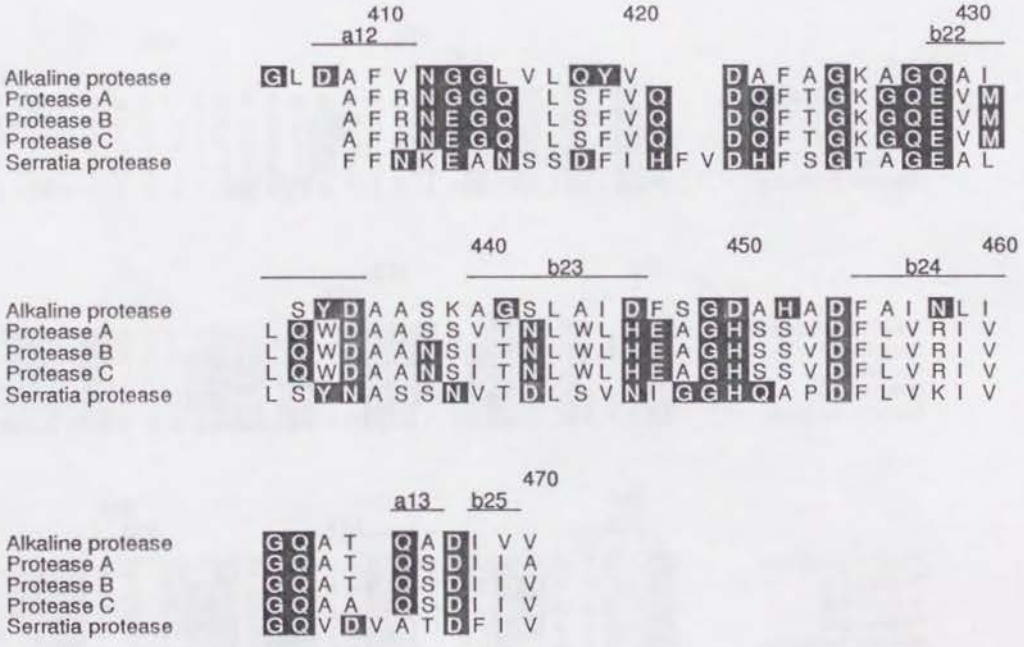


Figure 2.29: —*continued.*

2.4 The C-terminal domain

The C-terminal domain comprises residues 251–470 with eight Ca^{2+} ions (Ca1–Ca8 in Figure 2.25). The domain is highly β-strand dominant, containing eighteen β-strands, β8 to β25. Its approximate size is 47×25×21 Å, exhibiting an elongated shape.

The β-helix region

The β-strands in the domain are incorporated in a two-layer β-sandwich structure as shown in Figure 2.30. The domain can be subdivided into three separate regions on the basis of structural features. The proximal region to the catalytic domain, which includes five strands, β8 to β12, adopts a mixed parallel/antiparallel β-sheet structure with irregularly wound loops. The region is connected to the proteolytic domain by a loop linking α9 to β8, and is in contact with the proteolytic domain through one external face of this region. The C-terminal region consisting of strands β20 to β25 also has a mixed β-sheet topology with rather irregular loop connections between the successive strands. The central region consisting of six strands, β14 to β19, has a right-handed parallel β-helix (RβH) structure formed by all the strands arranged regularly in a spiral.

The N-terminal α-helix, α1, is packed predominantly on one external face of the β-sheet (β15, β17, β19) of this region. In the right-handed parallel β-helix (RβH) structure, two parallel β-sheets, (β15, β17, β19) and (β14, β16, β18), are packed together in an antiparallel manner. The sheets of the central β-helix have only a slight twist, which is in striking contrast to parallel β-sheets found in parallel α/β domains of proteins. This flatness of the sheets in the β-helix seems to be a consequence of the characteristic Ca^{2+} binding discussed below. This is also in contrast to the nature of the β-sheets in the other regions of the β-sandwich structure, which are more individually twisted. The inside of the β-helix can be further subdivided into three regions on the basis of their water molecule accessibility. The central core region of the β-helix is kept highly hydrophobic, because each set of Leu349, Leu367, Leu392, and Leu340, Leu358, Phe376 makes parallel lines stacked like a zipper by hydrophobic contact, as shown in Figure 2.31. As a result, there is no water molecule in this region. On the other hand, 'upper' and 'lower' regions from the central core region are hydrophilic because four aspartate residues, Asp338, Asp356, Asp374 and Asp400, are arranged in a straight line in the upper region to form the coordinate bonds with Ca3, Ca4 and Ca5. Asn347, Asp365 and Asp390 also line up in the lower region and form coordinate bonds with Ca6 and Ca7, as shown in Figure 2.31.

Parallel right-handed β-helix structures (RβH) similar to the central RβH are also found in pectate lyases C and E from *Erwinia chrysanthemi* [96, 97],

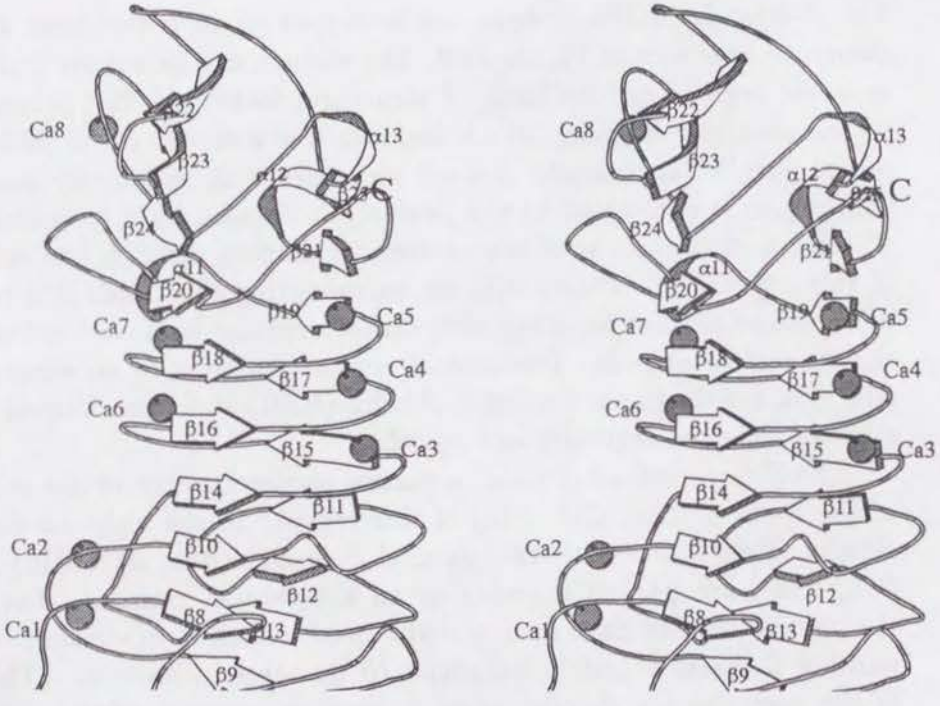


Figure 2.30: C-terminal domain of the alkaline protease. The domain comprises residue 251–470 with eight Ca^{2+} ions (Ca1–Ca8). The domain is predominantly formed by β-strands, β8–β25, β14–β19 of which form a parallel right-handed β-helix (RβH) structure. The domain can be further subdivided into three subdomains: catalytic domain proximal region (β8 to β13), C-terminal region (β20 to β25), and β-helix region (β14 to β19).

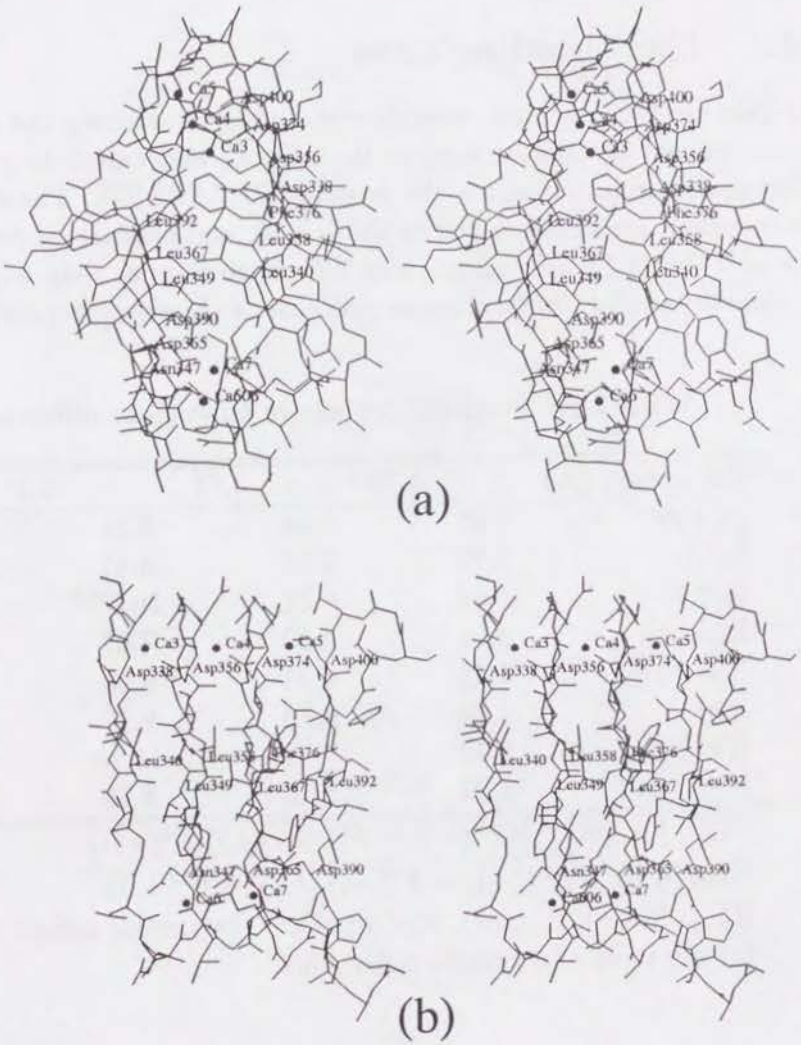


Figure 2.31: Stereo drawings of skeletal models in the parallel right-handed βhelix (RβH). (a) A stereo drawing of the β-helix showing the inside. Two lines of three leucine residues, each of which comprises three of the residues, make a hydrophobic central region stacked together. The upper side of the central region, the 'upper' region, contains three calcium ions, Ca3, Ca4 and Ca5. The lower region contains two calcium ions, Ca6 and Ca7. (b) The same region of the β-helix obtained by rotating (a) by 90° about an axis running from the top to the bottom of the paper.

and the tailspike protein of *Salmonella typhimurium* phage P22 [98]. In all of the structures, the β-strands are connected by right-handed crossovers which seem to be essential for constructing the β-helix structure.

2.4.1 Ca^{2+} binding sites

The calcium binding sites were determined by measuring the anomalous dispersion effects of calcium ions in the crystal, and calculating an anomalous difference Fourier map using the program PHASES [69]. The statistics of this measurement are summarized in Table 2.16, and anomalous peak heights and sites in Table 2.17. An anomalous difference Fourier map superimposed on the skeletal model of the alkaline protease is shown in Figure 2.32.

Table 2.16: Statistics for native anomalous differences.

Resolution (Å)		4.76	3.78	3.30
R(+)[a)]	3.66	5.59	8.31	13.13
R(−)	3.79	5.54	8.47	12.83
R(±)	4.66	6.87	10.37	15.91
R(cent)	3.01	4.59	7.37	10.91
Rw(+)[b)]	4.31	6.34	9.08	13.98
Rw(−)	4.48	6.20	9.17	14.00
Rw(±)	5.42	7.68	11.18	17.06
Rw(cent)	4.28	5.53	8.78	11.61

[a)] $R(+) = 100 \sum |\langle F^2(+)\rangle - F^2(+,i)| / \sum F^2(+,i)$.

[b)] $Rw(+) = \sqrt{\langle F^2(+)\rangle - F^2(+,i)}^{\,2} / \sum wF^2(+,i)$

R(−), Rw(−), R(cent), Rw(cent) are the similar definitions for for negative and centric reflections.

The β-helix of the alkaline protease is stabilized through Ca^{2+} binding to the sites specialized with the Ca^{2+} binding motif, **GGXGXDXBX** (B: bulky hydrophobic residue, ideally leucine). In the motif, loops connecting each of the β-strands in the β-helix make sharp turns bending approximately 180°, which provide liganding basements for the Ca^{2+} ions. This motif is repeated four times in total, between the sequence Gly334–Val378.

The Ca5 was found to be replaced by an UO_2^{2+} ion when the native crystal was soaked into the UO_2^{2+} solution to prepare the heavy-atom derivative. Therefore, the binding site for Ca5 is inferred to be more susceptible to ligand exchanging reactions, maybe due to looser stereo hindrance around the site than inside the β-helix.

Table 2.17: Peak assignments in an anomalous difference Fourier map. The atom names in this table correspond to those of Figure 2.25. Fractional coordinates in the unit cell are shown for the peak sites. The eighth calcium ion, Ca8, could not be assigned in this peak search.

Peak number	Hight	X	Y	Z	Atom name
1	296	0.7834	0.2029	0.5134	Ca7
2	275	0.3177	0.2357	0.1747	Ca4
3	265	0.8283	0.2083	0.8626	Ca5
4	259	0.7946	0.2292	0.4893	Ca6
5	248	0.8286	0.2370	0.8481	Ca3
6	197	0.4465	0.1977	0.5758	Ca1
7	179	0.4028	0.2211	0.6024	Ca2
8	170	0.0575	0.1228	0.5641	-
9	168	0.2958	0.0943	0.7739	-
10	167	0.4527	0.1811	0.3779	-
11	166	0.3413	0.2126	0.8268	-
12	164	0.2645	0.0451	0.4997	-
13	164	0.3759	0.1734	0.6994	-
14	163	0.6396	0.0569	0.5506	-
15	161	0.3873	0.1598	0.2132	-
16	161	0.2550	0.2503	0.1995	-
17	161	0.4309	0.1179	0.6624	-
18	161	0.7680	0.1455	0.7382	-
19	160	0.7329	0.2388	0.6099	-
20	159	0.1489	0.2143	0.2139	-

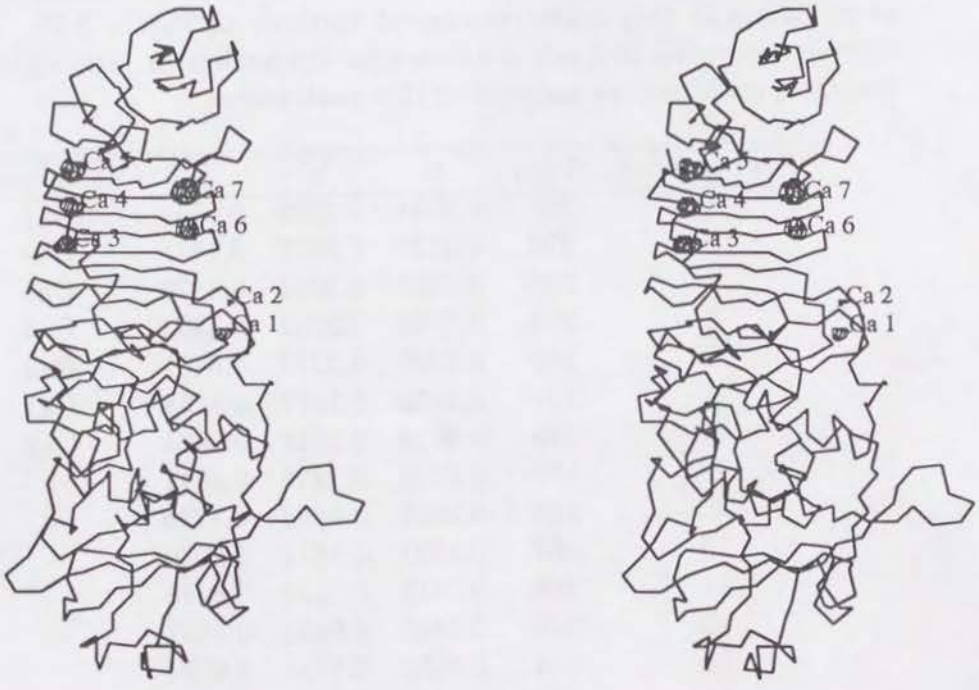


Figure 2.32: A stereo drawing of a native anomalous difference Fourier map at 3.0 Å resolution. The map, contoured at 4σ level, is superimposed on the Cα model of the alkaline protease. A calcium ion Ca8 could not be directly identified using this map, but could be assigned to an equivalent peak on the difference Fourier and $2|F_{\mathbf{obs}}| - |F_{\mathbf{calc}}|$ maps during the refinement procedure.

2.4.2 Calcium coordination modes

The modes of Ca^{2+} coordination are summarized in Figure 2.34. Each of the calcium ions adopts six-coordination with the main chain carboxyl oxygens (O), side chain oxygen (OD1, OD2, OE1, OE2) and water molecules (H_2O), except for Ca2 which adopts seven-coordination.

A stereo drawing of the portion around Ca8 is shown in Figure 2.33. The

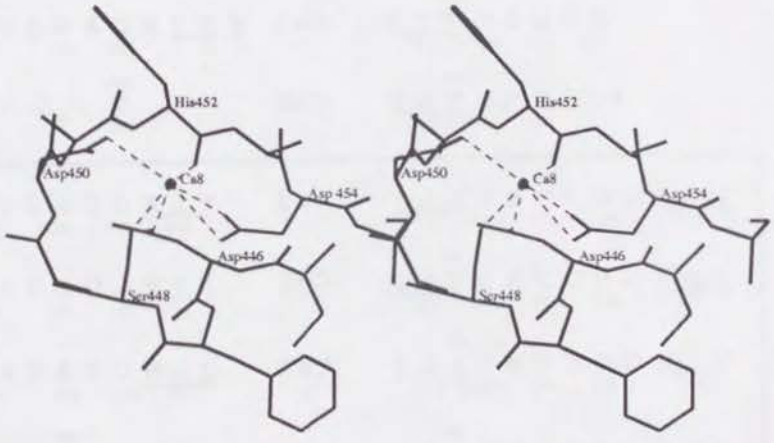


Figure 2.33: Coordination geometry of Ca8. The ion has six ligands, OD1 and O of D446, OD1 of Asp450, O of His452, OD1 and OD2 of Asp454.

Ca8 ion is a distinctive calcium ion compared with the case of *Serratia* protease which has seven calcium ions corresponding to Ca1–Ca7 of the alkaline protease. The circle-like loop region comprising residues Phe447–Asp454 between β23 and β24 makes a calcium binding sites for Ca8. The temperature factor for Ca8 shows an abnormally high value (72.9 Å^2), probably because of low occupancy of the ion[12]. Therefore the Ca8 is thought to be less significant for the structural stability of the alkaline protease than the other calcium ions.

2.5 Discussion

2.5.1 Structural comparison with other metzincins

Zinc metalloendoproteases hitherto studied are classified into five families as shown in Figure 2.1, page 10. All of them possess the zinc atom at their

[12]The occupancy of Ca8 was assumed to 1 during the refinement processes.

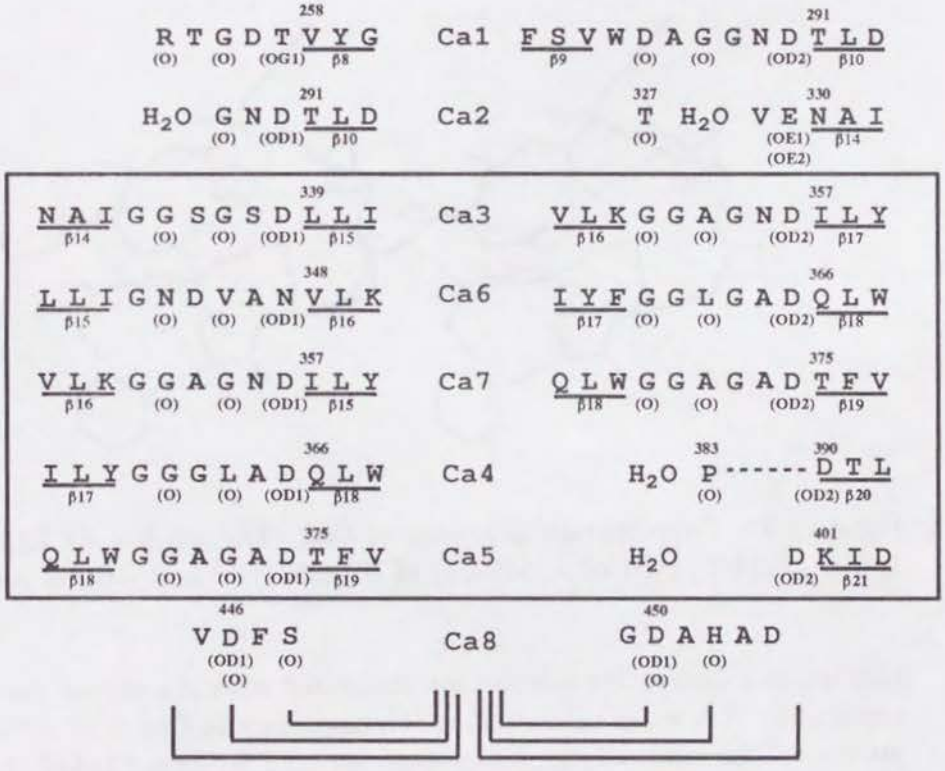


Figure 2.34: Ca^{2+} coordination modes. The boxed region indicates residues and ligands in the inside of the β-helix. All of the calcium ions adopt six-coordination geometries, except for Ca2, having a seven-coordination. The repeat of the motif, "GGXGXDXBX" is observed in the boxed region.

catalytic centers, where the zinc binding motif, **HEXXH**, is present as in the cases of bacterial endopeptidase, thermolysin, depicted in Figure 2.36, and closely related neutral endopeptidases [43, 40, 99, 100]. In the motif, two histidine residues serve as zinc ligands, and the glutamic carboxylate is thought to polarize a zinc-bound water molecule for a nucleophilic attack on the scissile peptide bond of a bound substrate [101], as illustrated in Figure 2.2, page 12. The fourth ligand to the zinc atom is present 20 residues downstream from the second histidine ligand, which is in another sequence motif **NEXXSD** embedded in an α-helix running perpendicular to the active site helix. However, the **HEXXH** motif has been observed not only in thermolysins, but also in other various kinds of metalloendoproteases [102]. Some of these contain the further elongated zinc binding motif, **HEXXHXXGXXH** [39]. They are grouped into the superfamily 'metzincins' because all of them have a strictly conserved 1,4-β turn, shown in Figure 2.1, containing a methionine residue beneath the catalytic zinc atom. A sequence alignment of the metzincins based on their tertiary structures was carried out to reveal similarities in their secondary structure elements, as shown in Figure 2.37 The conserved methionine is believed to offer a hydrophobic basement for their zinc ligands, beneath the catalytic zinc [39].

The metzincin superfamily consists of four families—serralysins, astacins, matrilysins and adamalysins. The serralysins are named after the representative enzyme, *Serratia* protease. The primary structure of the enzyme was deduced from the gene sequence of a gram-negative bacterium *Serratia marcescens* [103], and its tertiary structure has recently been determined at atomic resolution [104]. Tertiary structures of metzincins have been studied using X-ray crystallography, and are now available at atomic resolution. To reveal folding similarities among metzincins and significant structural differences which may be associated with functions, astacin [40, 105], a catalytic domain of human fibrobrast collagenase complexed with a peptide hydroxamate inhibitor [106, 107], adamalysin II [108], and a ligand-free form of *Serratia* protease [104] were each superimposed on the ligand-free alkaline protease using the program LSQKAB (CCP4 program suit), as shown in (a), (b) and (c) of Figure 2.38. A thermolysin complexed with a phosphoramidon inhibitor [47] was also superimposed on the alkaline protease. The folding topologies of metzincins turned out to be almost same, in that three α-helices, α5, α6 (active helix) and α9, and a β-pleated sheet comprising five β-strands, β-3 to β-7, are present in the 'upper' domain. As shown in Figure 2.37 and Figure 2.38, these secondary structure elements and Met-turn are aligned in the order of β3, α5, β4, β5, β6, β7, α6 (active helix), 'Met-turn', and α9.

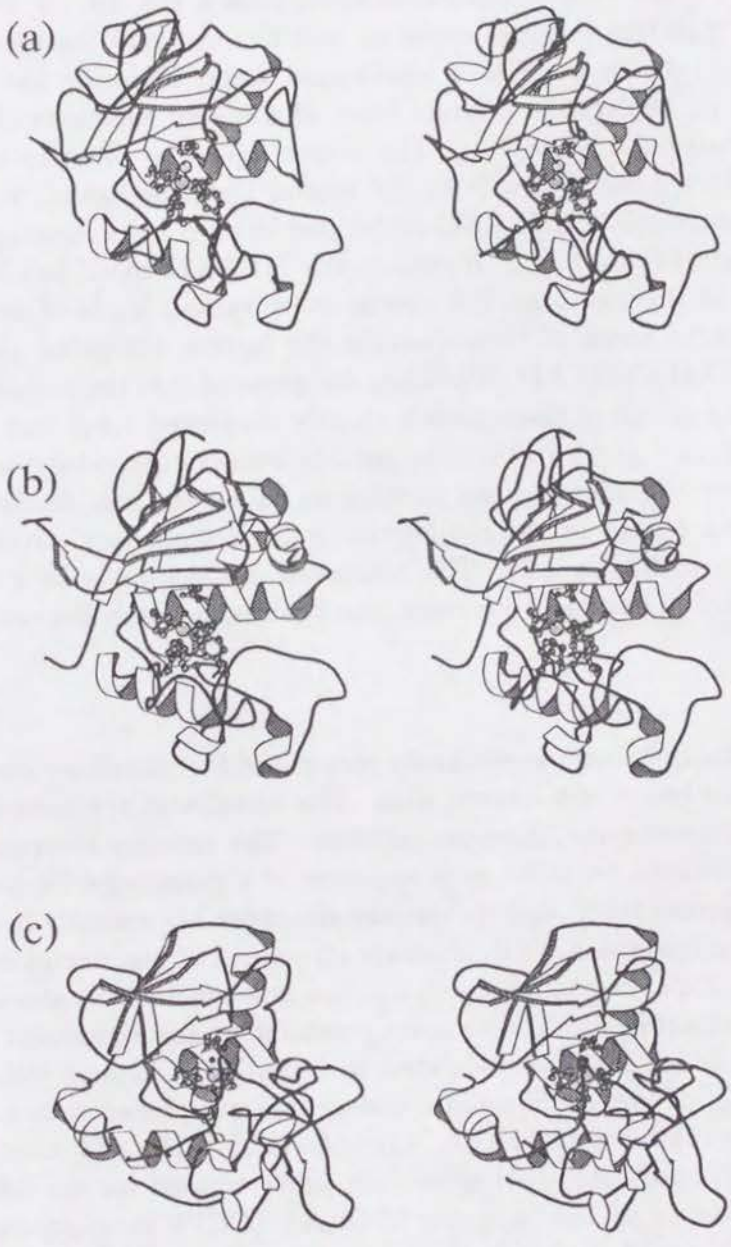


Figure 2.35: Tertiary structures of zincins. (a) The catalytic domain of the alkaline protease, (b) the catalytic domain of the *Serratia* protease, (c) Astacin, (d) the catalytic domain of human fibrobrast collagenase with an inhibitor, (e) adamalysin, and (f) thermolysin with an inhibitor.

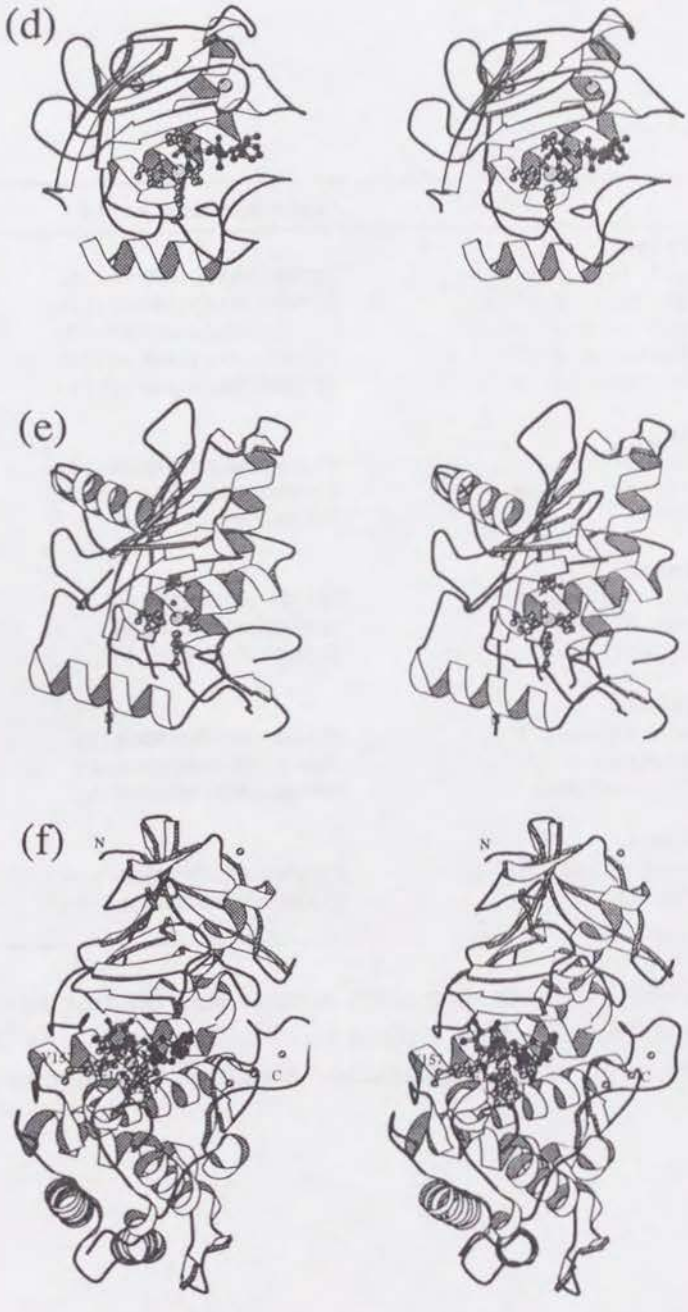


Figure 2.35: —*continued.*

	zinc binding motif	Met-turn
Serralysins		
alkaline protease	TLTHEIGHALGLSHPGD(16)	ESTRAYSVMSYWEEQN
serratia protease	TFTHEIGHALGLSHPGD(16)	EDTRQFSLMSYWSETN
protease A	TLTHEIGHALGLSHPAE(17)	EDSRQFSIMSYWSIEN
protease B	TLTHEIGHALGLAHPHE(16)	EDSYQFSIMSYWDIEN
protease C	TFTHEIGHALGLAHPGE(16)	EDSYQFSIMSYWGENE
Astacins		
astacin	TIIHELMHAIGFYHEHT(33)	EDYQYYSIMHYGKYSF
meprin (mouse)	TIEHEIFHALGFFHEQS(35)	TPYDYESLMHYGPFSF
BMP1	IVVHELGHVVGFWHEHT	----IGALMHYARNTF
Matrilysins		
collanenase f	VAAHELGHSLGLSHSTD	----IGALMYPSYTFS
collagenase n	VAAHEFGHSLGLAHSSD	----PGALMYPNYAFR
matrilysin	AATHEHELGHSLGMGHSS	----PNAVMYPTYGNG
Adamalysins		
adamalysin II	TMAHELGHNLGMEHDGK(2)	RGASL-CEMRPGLTPG
atrolysin	TMAHELGHNLGMEHDGK(2)	RGASL-CEMRPGLTKG
trimerelysin	TMTHELGHNLGMEHDDK(3)	CEA---CEMSDVISDK
Thermolysin		
thermolysin	VVAHELTHAVTDYTAGL----	
elastase	VAAHEVSHGFTEQNSGL----	

Figure 2.36: Sequence alignment around zinc binding sites of five zinc metalloendoprotease families around zinc binding sites. The 'Met-turn' motif is conserved in all of the metzincins. A part of the alignment was cited from [39].

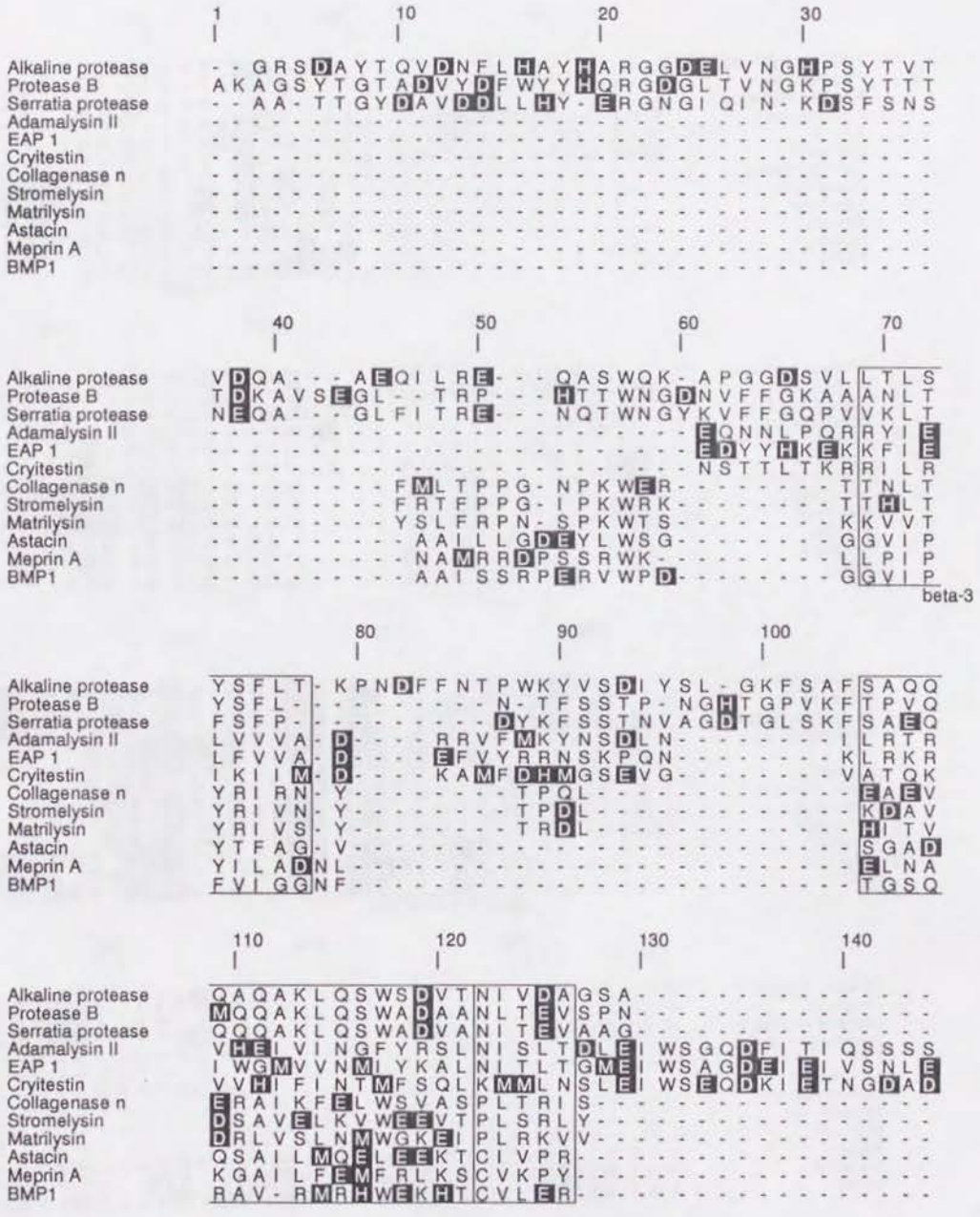
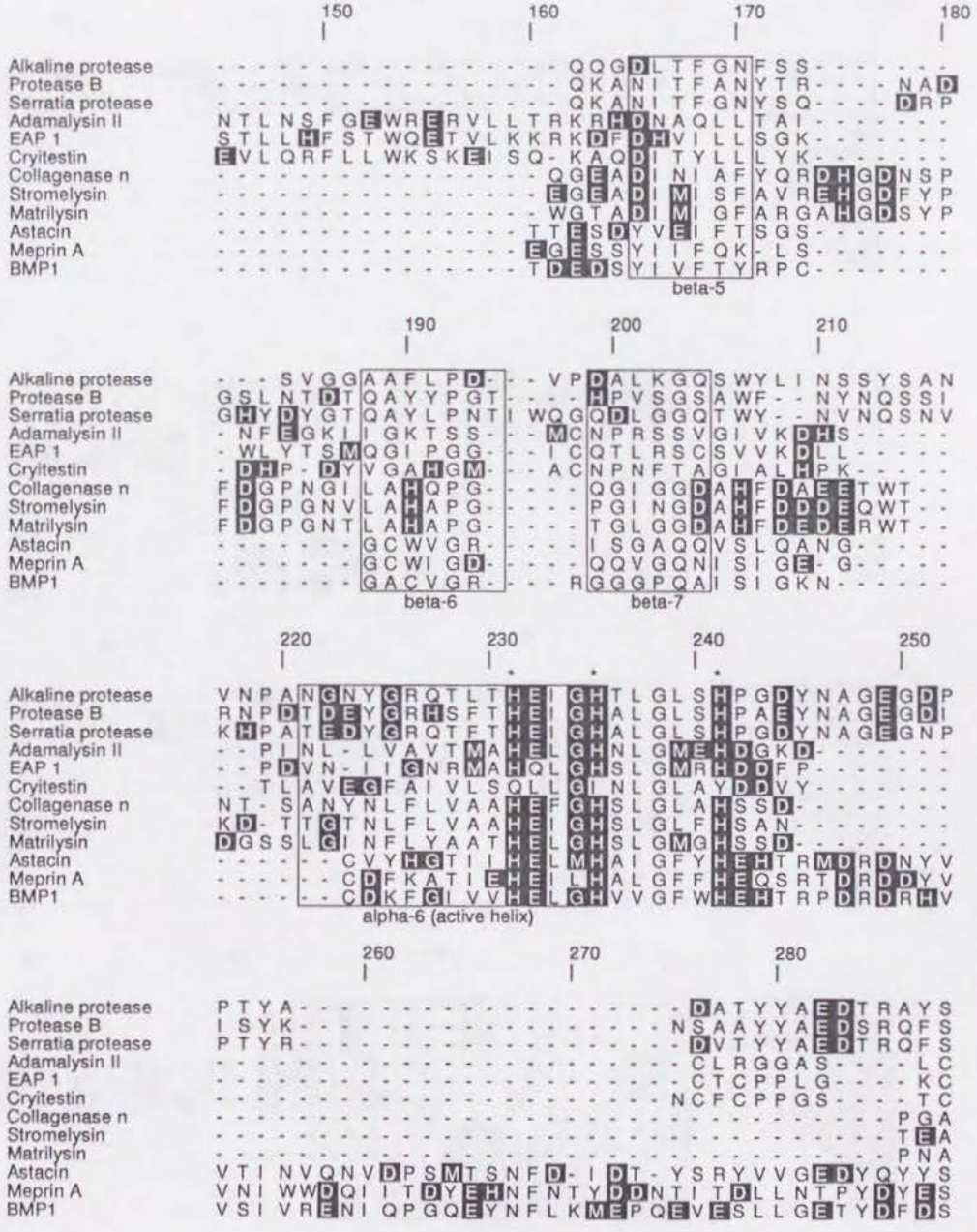


Figure 2.37: Sequence alignment of metzincins.


150
160
170
180
Alkaline protease
Protease B
Serratia protease
Adamalysin II
EAP 1
Cryitestin
Collagenase n
Stromelysin
Matrilysin
Astacin
Meprin A
BMP1
beta-5
190
200
210
beta-6
beta-7
220
230
240
250
alpha-6 (active helix)
260
270
280

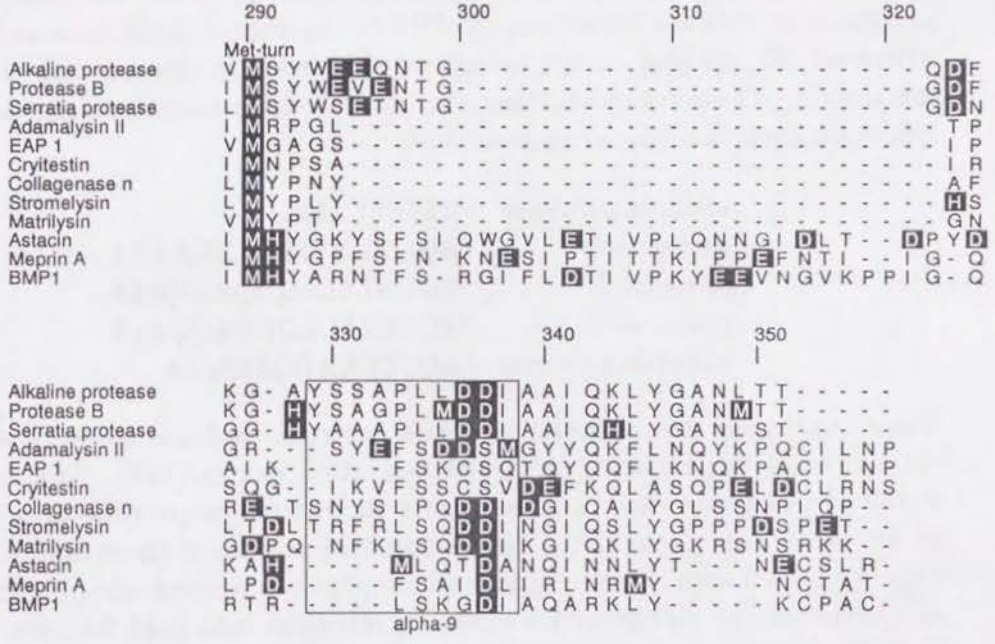


Figure 2.37: —*continued.* **Serralysins**: alkaline protease, protease B and *Serratia* protease. **Adamalysins**: adamalysin II, EAP 1 and cryitestin. **Matrilysins**: collagenase n, stromelysin and matrilysin. **Astacins**: astacin, meprin A and BMP1. Designations for secondary structures correspond with those in Figure 2.26. Part of this alignment was cited from [38].

N-terminal extensions

As shown in Figure 2.37, the serralysins (alkaline protease, *Serratia* protease and protease B) have the extension of 40–50 residues in the N-terminal region, where $\alpha 1$, $\beta 1$, $\beta 2$ and $\alpha 2$ are present in this order, compared with the other metzincins. These regions further have short N-terminal extensions, as shown below, in their inactive zymogens [109].

alkaline protease	MSSNSLALK
protease A	MQQNENVSLNTSAAAKA
protease B	MGKNLSLRQDDAQQALS
protease C	MGKNLSLRQDDAQHALS
Serratia protease	MQSTKKAIEITESSLA

These zymogens are secreted into the external medium where a divalent cation (most likely Ca^{2+}) mediated activation occurs [109]. The short N-terminal extensions are removed during the activation process, probably by an autocatalytic mechanism. The amino acid profiles of these extensions do not correspond with those of typical signal peptides where hydrophobic amino acid residues are dominant for enabling intrusion into lipid bilayers. Questions therefore arise as to what functions these short N-terminal extensions possess. Taking into account detailed studies on the secretion machinery of proteases B and C which are highly homologous to the alkaline protease, these extensions may be essential not for secretion, but rather for rendering the protease inactive. In the inactive forms, the N-terminal portions may adopt different conformations from those of native forms so that the short N-terminal extensions interact with the active site and then are cleaved by autolysis. The N-terminal extensions might also be essential for the proper folding of the mature protease, as has been shown for subtilisin where the pre-pro-subtilisin transformation occurs [110]. Moreover, in all the families except for adamalysins, a conserved Trp residue (designated W57 in Figure 2.37) is located in a superimposable bend between $\beta 6$ and $\beta 7$ of the pleated sheet. The tryptophan residue is buried in a hydrophobic environment, which might imply that the conserved residue plays a significant role in folding into mature forms.

The upper bulge region of the alkaline protease

One of the most remarkable structural differences between the alkaline protease and other metzincins is the inserted region between $\beta 3$ and $\alpha 5$, shown in Figure 2.37. As a result of the insertion, this region, approximately comprising P65–F88, forms an 'upper bulge region' over the catalytic cleft of the alkaline protease. As discussed later, this bulge turned out to be associated with inhibitor binding. The side chain of Asp78 makes salt bridge with

$N\eta_{1,2}$-Arg of the inhibitor, which corresponds with the observation that the arginine residue in the peptide portion of the inhibitor enhances the affinity to the enzyme [65, 66]. However, the bulge does not seem to be associated with substrate binding.

In the case of adamalysins, a long inserted portion consisting of about 30 residues is located between β4 and β5. This portion folds into an additional α-helix over the five-stranded β-pleated sheet, making a projection which corresponds to the upper bulge of the alkaline protease. However, it may not be directly associated with any functional aspect of substrate binding, since it is not elongated over the catalytic cleft, which means that the projection cannot interact with the substrate.

With the exception of astacins, the portion between β5 and β6 makes another bulge region. In the alkaline protease, this region comprises four residues, Ser-Val-Gly-Gly, and bulges out over the catalytic cleft almost perpendicular to β5 and β6, as shown in Figure 2.38. This bulge spatially restricts the ligand binding to the enzyme. In contrast, the corresponding region of adamalysin II is displaced to the right approximately by 6 Å, even though the region also makes a bulge (Figure 2.38 (c)). In a collagenase, this region does not make a bulge, but forms a part of the rim of the upper wall along the active site cleft. As a result, the rim of the upper wall of the collagenase is longer than that of the alkaline protease (Figure 2.38 (b)). Even in the *Serratia* protease, this region has a 'more open' conformation than that of the alkaline protease. However, the bulges or the rims of this region play a certain role in restricting substrate binding. In contrast, the astacin has no bulge or rim in this region, making it the least spacially restricted area. This might correspond with the observation that the astacin possesses multiple substrate binding sites, and only hydrolyzes pentamers or longer substrates [105, 111].

The region comprising β5–α6 (active helix) has relatively less insertions or deletions than other regions. This implies that these are functionally important. Actually, as can been seen in residues around α6, catalytically invaluable residues, comprising three histidines and one glutamic acid, are located in this region making a zinc binding site. The serralysins have short insertions containing a portion Tyr-Asn-Ala-Gly-Glu-Gly, which makes a flexible loop region between α6 and the Met-turn. This flexible loop region turned out to be extremely significant for the serralysins to bind and stabilize the substrates, as discussed later. In short, the conformation of the flexible region changes on ligand binding, leading to the formation of hydrogen bonds with the ligands, and with the phenolic oxygen of Tyr216 to stabilize the Michaelis-complexes. In the case of the astacin, a longer region is inserted between α6 and the Met-turn, leading to the formation of the largest 'lower' wall found among the metzincins as seen in Figure 2.38 (a), as expected from the longest insertion between α6 and the Met-turn. Atoms on the lower wall

of the astacin , however, has relatively lower temperature factors than those of the alkaline protease, leading to the conclusion that the lower wall of the astacin takes a rigid conformation during hydrolysis, which contrasts with the catalytic mechanism of the alkaline protease discussed later.

Met-turns and S_1' pockets

The $\beta 4$ is followed by the strictly conserved 'Met-turn' as shown in Figure 2.38. The Met-turn is located beneath the active site zinc ion, and forms the hydrophobic environment for zinc ligands. In serralysins and astacins, the conserved methionine residue in the Met-turn is followed by the tyrosine residue which serves as the fifth zinc ligand. In contrast, adamalysins and matrilysins possess no fifth zinc ligand; instead, bulge-like rims of the so-called 'S_1'' pocket are formed behind the Met-turn.

By accumulated evidence derived from a vast number of inhibitor-enzyme complexes, it has been shown that the S_1' pocket, which harbors the P_1' side chain of the substrate, plays a crucial role in substrate recognition and is the most influential determinant for enzyme specificity. As seen in Figure 2.38, all of the families except for the serralysin, have rims or walls of the S_1' pockets, which are designated 'S_1'' in the figure. Adamalysins and matrilysins cannot have a tyrosine residue as the fifth zinc ligand due to steric hindrance caused by the rims of the S_1' pocket. In the astacins, however, the rims are formed in a somewhat different manner from those of the adamalysins and the matrilysins. The S_1' rim of the astacins are built along the active site helix, juxtaposed to the catalytic zinc, thus providing a room for the fifth zinc ligand. In contrast, the region followed by Tyr216 (the fifth ligand of the zinc) makes a loop directed downward, causing the serralysins to have no S_1' rims or walls. As a result, the P_1' side chain of the substrate is less sterically hindered in the serralysins than in the other metzincins, implying that the serralysins have the relatively broad substrate specificity. As seen in Figure 2.5, page 16, the alkaline protease hydrolyzes the oxidized insulin-B chain in an almost non-specific manner, which seems to correspond with the structural observations made around the S_1' pocket.

2.5.2 Functions of the C-terminal domain of the alkaline protease

The catalytic (N-terminal) domains of metzincins have the basically common folding topology as discussed above. On the other hand, most of the metzincins have additional C-terminal domains, which are different from on another in their structures and functions. The C-terminal domains act as 'additional modules', taking part in a variety of functions. For instance, human nutrophil and fiblobrast collagenases have C-terminal domains which play

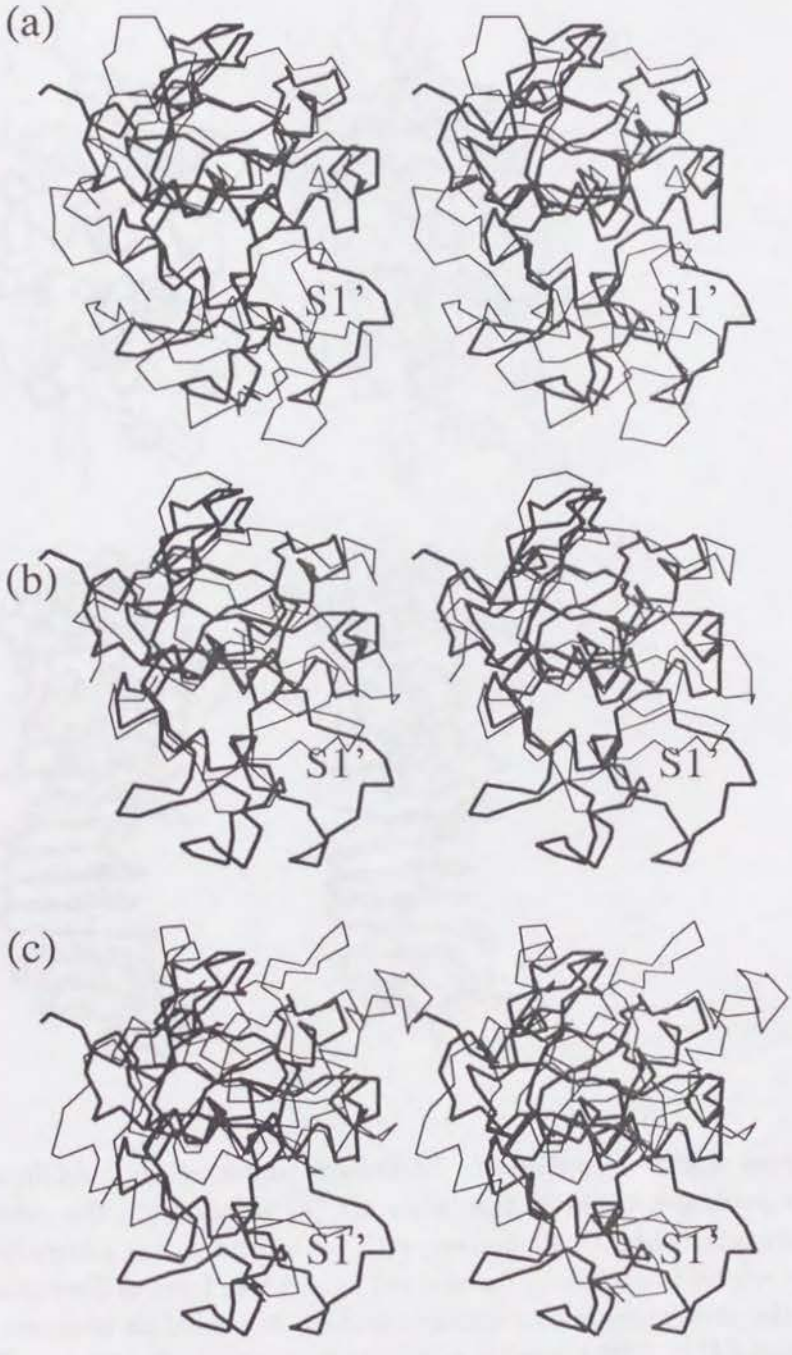


Figure 2.38: Structural comparison of zincins.

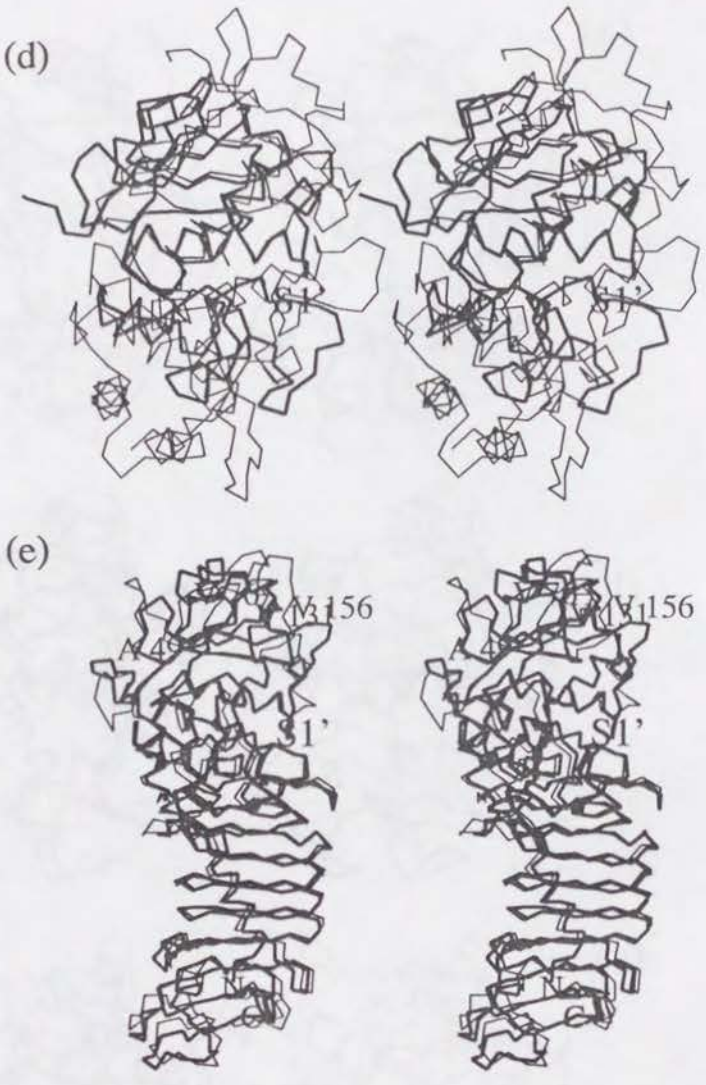


Figure 2.38: —*continued.* Structures depicted in bold lines are the alkaline protease, those in thin lines are (a) astacin, (b) the catalytic domain of human fibrobrast collagenase with an inhibitor, (c) adamalysin II, (d) thermolysin with an inhibitor, and (e) ligand-free form of *Serratia* protease. Each of the structures was superimposed on the alkaline protease using program LSQKAB (CCP4 program suit) and by manual adjustment. The sites around the S_1' pocket are signatured.

a role in unfolding triple-helical collagens. When the C-terminal domains are separated from the catalytic domains, their enzymatic activities are lost [112, 113, 114, 115]. The C-terminal domain of matrix metalloprotease (MT-MMP), which is a member of the matrilysins, is a functional transmembrane linker when it is fused to a secretory protein, and is a tissue inhibitor for matrix metalloprotease-1 [116]. The zymogens of the MT-MMPs are activated on the cell surface by the plasma membrane-dependent mechanism specific for the enzymes [117]. Furthermore, breast cancer cells have binding receptors on their surfaces for MMP-2, also a member of the matrilysins, and the bound MMP-2 can efficiently cleave surrounding matrix components [118]. Meprins A and B, which belong to the astacin family, also have the C-terminal domains which act as membrane linkers [119, 120]. In this context, in case of the alkaline protease, the C-terminal domain containing the β-helix region seems to be associated with a functional aspect.

Accumulated evidence suggests that the domain plays a important role in secreting the inactive zymogen of the enzyme into the external medium where the divalent cation-mediated activation occurs [109]. As shown in Figure 2.39, *aprA* (50.4 kDa), which is the structural gene of the alkaline protease, is located in a 8.8 kb fragment of *Pseudomonas aeruginosa* DNA linked with other gene open reading frames, *aprD*, (63.7 kDa) aprE (48.1 kDa), *aprF* (54.2 kDa) and *aprI* (13.9 kDa) [121]. The hydrophathy profile of the amino

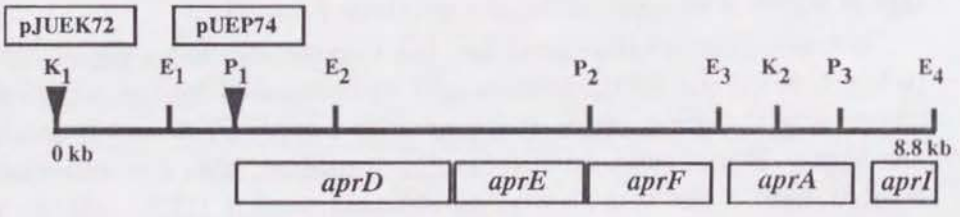


Figure 2.39: Genetic organizatin and restriction map of the *apr* locus derived from pJUEK72 [122]. Restriction sites indicated are *Kpn*I (K), *Eco*RI and *vu*II. The sites at which the lac promoter was inserted in the different plasmids are shown as solid downward arrowheads. Coding regions of *apr* genes are shown as open boxes [121]. The products encoded by the genes *aprD*, *aprE* and *aprF* together make a one-step secretion machinary for the alkaline protease. The *aprA* is a structural gene, encoding the zymogen of the alkaline protease, and the *aprI*, a macromolecular inhibitor for the enzyme.

acid sequence indicates that AprD lacks a signal peptide but possesses four hydrophobic domains corresponding to putative transmembrane segments, characteristic of inner membrane proteins. A motif for the ATP-binding site, **GXXGXGKS**, is located in the C-terminal region of AprD. The hydropathy

profile of the second gene, AprE, shows one hydrophobic domain located between residues 14 and 36, without signal peptide. The third, AprF, is largely hydrophilic with the exception of residues at the N terminus which show properties of a typical signal sequence. As shown by the hydrophathy profiles of amino acid sequences encoded by the genes, the products encoded by *aprD*, *aprE* and *aprF* may together make the special secretion machinery driven by the hydrolysis of ATP.

Many Gram-positive and Gram-negative bacteria secrete extracellular proteins into the growth medium [80]. In the case of Gram-positive bacteria, which only have a cytoplasmic membrane, most of them are secreted via the "signal peptide" pathway. In the case of Gram-negative bacteria including *Pseudomonas aeruginosa*, in which the cell envelope comprises two membranes, the secretion of extracellular proteins always involves a specific mechanism. For some of the proteins secreted by Gram-negative bacteria, the initial step involves the recognition of a typical N-terminal signal peptide by the general export system and the subsequent translocation of the protein across the cytoplasmic membrane, followed by another step requiring specific functions to cross the outer membrane. There are also some secreted proteins which do not have a signal peptide. For instance α-hemolysin, the product of the *hlyA* gene in hemolytic *Escherichia coli* strains, does not have a signal peptide. Two envelope proteins, the products of the *hlyA*-linked genes *hlyB* and *hlyD*, are necessary for α-hemolysin secretion [80, 123, 124], and the vicinity of the C-terminal portion of α-hemolysin contains a signal region which is recognized by the envelope proteins.

In case of the alkaline protease, the enzyme also has a signal region close to the C-terminus [121], which might be recognized by the secretion system comprising the products of *aprD*, *aprE* and *aprF*. Proteases B and C, which are highly homologous to the alkaline protease, also has secretion signals located within the last C-terminal 40 amino acids [125], and the enzymes are secreted into an external medium via the similar secretion machinery. All these observations strongly suggest that the C-terminal domain of the alkaline protease has a signal region recognized by the one step secretion device consisting of the membrane proteins encoded by *aprA*-linked genes, *aprD*, *aprE* and *aprF*, as shown in Figure 2.39.

The product encoded by *apr I* is a macromolecule inhibitor for the alkaline protease. Recently, the complex between the *Serratia* protease and a macromolecule inhibitor from *Erwinia chrisanthemi* homologous to that of the alkaline protease has been solved by X-ray crystallography[126]. In this complex, the N-terminal portion of the inhibitor was located inside the catalytic cleft of the *Serratia protease*, which accommodates the active site zinc ion and general base residue Glu177 [13]. The *Erwinia chrisanthemi* inhibitor

[13]The picture of this complex cannot be drawn at present, because the set of coordinates

is fold in a compact eight-stranded antiparallel β-barrel of simple up-down topology as is found in members of the retinol binding protein family, and occupies the S_1' site of the *Serratia* protease. The product of the gene *apr I* probably inhibits the zymogen of the alkaline protease in the same manner as that of the *Serratia* protease–*Erwinia chrisanthemi* inhibitor complex, and prevents the components inside the bacterium from uncontrolled hydrolysis caused by the enzyme. As mentioned above, the C-terminal domain of the alkaline protease consists of three separate subdomains on the basis of structural features: the catalytic domain proximal region, the C-terminal region, and the β-helix region. Through sequence similarities with the proteases B and C, the C-terminal region of the alkaline protease, which comprises strands β20 to β25, may be recognized as a signal region by the secretion machinery encoded by *aprD, aprE* and *aprF*. The C-terminal region may be constructed immediately after the translation of the structural gene, *aprA*, offering a recognition region for the secretion machinery.

In contrast, the β-helix region is inferred to be unstable or almost unfolded in the cell, because the Ca^{2+} concentration inside of bactera is known to be relatively low compared with that of the outer medium, and the β-helix requires Ca^{2+} to fold stably. It is widely accepted that most of the systems for protein secretion or translocation across biological membranes require the following principale mechanisms [80];

1. The protein to be secreted across a membrane must adopt a "translocation competent" conformation, most likely an incompletely folded conformation. Special proteins seem to be necessary to promote its establishment or to increase its half-life.

2. The protein to be secreted must somehow interact with the membrane. In most cases, this is achieved by a signal peptide which is subsequently cleaved off.

3. In cases where two membranes have to be crossed (chloroplast envelope, mitochondrial outer and inner membranes, and envelope of Gram-negative bacteria such as: *Pseudomonas aeruginosa, Serratia marcescens*, and *E. coli*) contact sites between the two membranes have been implicated in the process.

4. Biosynthesized proteases must be inactive in the cell of the bacterium to avoid hydrolysis harmful to the organisms.

In the case of the alkaline protease, the unfolded β-helix region seems to enable the enzyme to adopt the "translocation competent" conformation.

---

designated '1SMP is now hold on in PDB. Refer figure 1–3 in [127].

Very interestingly, α-hemolysin also has eleven tandem repeats of the 'GGXGXD' motif between the amino acid residue, 789 and 849 [128], and the higher molecular weight zymogen of the α-hemolysin is activated by Ca^{2+} after secretion into the extracellular medium [123]. The **GGXGXD** repeated region may fold into the right-handed β-helix (RβH) topology as in the cases of the alkaline protease [129, 130, 127] and *Serratia* protease [104].

Taking into account all of these observations, the roles of the C-terminal domain containing the β-helix region may be deduced as follows;

1. The alkaline protease lacks a typical signal sequence in the N-terminal region. This implies that the enzyme requires a special secretion system which is independent from the general secretion systems.

2. The special secretion system consists of the membrane proteins encoded by genes, *aprD*, *aprE*, and *aprF*, which are located in the vicinity of the structural gene, *aprA*, as shown in Figure 2.39. This secretion system is driven by the hydrolysis of ATP.

3. The β-helix region of the alkaline protease may adopt an almost unfolded conformation suitable for passing through the membrane of the bacterium.

4. The macromolecule inhibitor of the enzyme, encoded by *aprI* located proximally to *aprF*, must inhibit the catalytic activity of the enzyme in the inside of the bacterium to avoid harmful hydrolysis by the enzyme.

5. The C-terminal region containing strands β20 to β25 may be recognized by the secretion mechanism. The macromolecule inhibitor might be separated from the enzyme before the enzyme passing through the membrane by using the secretion mechanism.

6. The higher-molecular-weight zymogen of the alkaline protease is subsequently activated by Ca^{2+} so that the N-terminal extension is cleaved off. In this step, the unfolded β-helix region folds into a right-handed helical conformation, and then the enzyme autocatalytically hydrolyzes the N-terminal extension.

Chapter 3

Crystal structures of complexes

3.1 Introduction

As mentioned previously, one of the main purposes of this study is to reveal the catalytic mechanism of the alkaline protease. The detailed elucidation of the mechanism will solve the questions raised previously (*c.f.* Figure 2.6), whether the alkaline protease is specific to a substrate which has a lysine or an arginine in its P_1 site [64, 65, 66]. It is difficult, however, to understand the mechanism by solving only the native structure, because X-ray structures are usually steric and do not shed much light on the dynamics of catalysis. Although dynamic crystal structures can be obtained using a time-resolved Laue method, the method is still under development and has some difficulties to be solved [131, 132]. Therefore, complexes are usually prepared in order to reveal residues, water molecules, metal ligands and conformational changes which during the catalytic process. In general, there are two ways to prepare crystals complexed with ligands, *e.g.*, inhibitors, substrate analogues, transition state analogues and macromolecules. They are;

- The soaking method where are soaked native cystals into solutions of the ligands for preparing crystals of the complexes. attempted into introduce to the crystals.
- The co-crystallization method where enzymes are complexed with ligands in solution *prior* to being crystallized.

In earlier studies on catalytic mechanisms of thermolysin [46, 47, 42, 48, 45, 44] or carboxypeptidase A [19, 133, 134, 135], almost all the complexes were prepared by the soaking method. In the recent studies, some complexes of the collagenases [136, 107] and *Serratia* protease [104], which are members of the metzincins (*c.f.* Table 2.1, page 10) have been reported to be prepared by the

co-crystallization technique. For preparing complexes between the alkaline protease and a specific inhibitor or substrate analogue, these two methods were tried. Many specific inhibitors for various metalloproteases have been developed with the common strategy of using the most susceptible peptides (*see* Figure 2.6, page 17) with appropriate ligand groups to zinc. A specific inhibitor was originally developed for the *Serratia* protease [66];

$$\text{Bz-Phe-Arg-(NH)-(}C_6H_4\text{)-}m\text{-S-Bzl·HCl (Mw = 736.8)}$$

This inhibitor was found to be also effective to the alkaline protease (K_i = 7.0 μM) [65]. Because it has been reported that some complexes of the metzincins were prepared by co-crystallization method [136, 107], the method was first tried to complex the alkaline protease with the specific inhibitor.

3.2 Materials and methods

Co-crystallization

A preparation of the enzyme-inhibitor complex by the co-crystallization method was conducted as follows;

1. 10.0 mg of the lyophilized alkaline protease was dissolved in 500 μl of a solution containing 1 mM of the inhibitor, 1 % (v/v) dimetylsulfoxide (DMSO), and 5 mM $CaCl_2$/1 mM NaN_3,

2. The protein solution was placed over stirring the solution in an incubator kept at 37 °C for 12 hours.[1]

3. After the incubation, the protein solution was filtered with the acetate filter (13 mm filter diameter, 45 μm pore diameter, Disposable Sterile Syringe Filter™; Corning Glass Works, Corning, New York, U. S. A.).

4. To search for crystallization conditions of the complex, the hanging drop vapor diffusion method was employed (*c.f.* Figure 2.7, page 19, using 97 kinds of reservoir solutions of a sparse matrix.)

5. 500 μl of each reservoir solution was poured into each well of the micro plate, one by one, and 2 μl of the reservoir solution poured was mixed with an equivalent volume of the complex solution on a siliconized glass cover slip.

6. The glass slips were set on the greased rims of the micro plates, and then the plates were placed at 25 °C or 4 °C.

Table 3.1: Screening conditions for co-crystallization using a sparce matrix. No.1–50 are from CRYSTAL SCREEN, and No.51–98, from CRYSTAL SCREEN II.

No.	Salt	Buffer	Precipitant
1.	0.02 M Ca chloride	0.1 M Na acetate pH 4.6	30 % v/v 2-methyl-2,4-pentanediol
2.	none	none	0.4 M K, Na Tartrate
3.	none	none	0.4 M NH_4 Phosphate
4.	none	0.1 M Tris HCl pH 8.5	2.0 M NH_4 Sulfate
5.	0.2 M Na citrate	0.1 M Na Hepes pH 7.5 2-methyl-2,4-pentanediol	30 % v/v
6.	0.2 M Mg chloride	0.1 M Tris HCl pH 8.5	30 % W/v PEG 4000
7.	none	0.1 M Na Cacodylate pH 6.5	1.4 M Na acetate
8.	0.2 M Na citrate	0.1 M Na Cacodylate pH 6.5	30 % v/v 2-propanol
9.	0.2 M NH_4 acetate	0.1 M Na citrate pH 5.6	30 % w/v PEG 4000
10.	0.2 M NH_4 acetate	0.1 M Na acetate pH 4.6	30 % w/v PEG 4000
11.	none	0.1 M Na acetate pH 4.6	1.0 NH_4 Phosphate
12.	0.2 M Mg chloride	0.1 M Na Hepes pH 7.5	30 % v/v 2-propanol
13.	0.2 M Na citrate	0.1 M Tris HCl pH 8.5	30 % v/v PEG 400
14.	0.2 M Ca chloride	0.1 M Na Hepes pH 7.5	28 % PEG 400
15.	0.2 M NH_4 Sulfate	0.1 M Na Cacodylate pH 6.5	30 % w/v PEG 8000
16.	none	0.1 M Na Hepes pH 7.5	1.5 M Li Sulfate
17.	0.2 M Li Sulfate	0.1 M Tris HCl pH 8.5	30 % PEG 4000
18.	0.2 M Mg acetate	0.1 M Na Cacodylate pH 6.5	20 % PEG 8000
19.	0.2 M NH_4 acetate	0.1 M Tris HCl pH 8.5	30 %v/v 2-propanol
20.	0.2M NH_4 Sulfate	0.1 M Na acetate pH 4.6	25 % w/v PEG 4000
21.	0.2 M Mg acetate	0.1 M Na Cacodylate pH 6.5	30 % v/v 2-methyl-2,4-pentanediol
22.	0.2 M Na acetate	0.1 M Tris HCl pH 8.5	30 % w/v PEG 4000
23.	0.2 M Mg chloride	0.1 M Na Hepes pH 7.5	30 % v/v PEG 400
24.	0.2 M Ca chloride	0.1 M Na acetate pH 4.6	20 % v/v 2-propanol
25.	none	0.1 M Imidazole pH 6.5	1.0 M Na acetate
26.	0.2 M NH_4 acetate	0.1 M Na citrate pH 5.6	30 % v/v 2-methyl-2,4-pentanediol
27.	0.2 M Na citrate	0.1 M Na Hepes pH 7.5	20 % v/v 2-propanol
28.	0.2 M Na acetate	0.1 M Na Cacodylate pH 6.5	30 % w/v PEG 8000
29.	none	0.1 M Na Hepes pH 7.5	0.8 M K,Na Tartrate
30.	0.2 M NH_4 Sulfate	none	30 % w/v PEG 8000
31.	0.2 M NH_4 Sulfate	none	30 % w/v PEG 4000
32.	none	none	2.0 M NH_4 Sulfate
33.	none	none	4.0 M Na Formate
34.	none	0.1 M Na acetate pH 4.6	2.0 M Na Formate
35.	none	0.1 M Na Hepes pH 7.5	1.6 M Na,K Phosphate

Table 3.1: —*continued.*

36.	none	0.1 M Tris HCl pH 8.5	8 % w/v PEG 8000
37.	none	0.1 M Na acetate pH 4.6	8 % PEG 4000
38.	none	0.1 M Na Hepes pH 7.5	1.4 M Na citrate
39.	none	0.1 M Na Hepes pH 7.5	2 % v/v PEG 400 2.0 M NH_4 Sulfate
40.	none	0.1 M Na citrate pH 5.6	20 % v/v 20 % w/v PEG 4000
41.	none	0.1 M Na Hepes pH 7.5	10 % v/v 2-propanol 20 % w/v PEG 4000
42.	0.05 M K Phosphate	none	20 % w/v PEG 8000
43.	none	none	30 % w/v PEG 1450
44.	none	none	0.2 M Mg Formate
45.	0.2 M Zn acetate	0.1 M Na Cacodylate pH 6.5	18 % w/v PEG 8000
46.	0.2 M Ca acetate	0.1 M Na Cacodylate pH 6.5	18 % w/v PEG 8000
47.	none	0.1 M Na acetate pH 4.6	2.0 M NH_4 Sulfate
48.	none	0.1 M Tris HCl pH 8.5	2.0 M NH_4 Phosphate
49.	1.0 M Li Sulfate	none	2 % w/v PEG 8000
50.	0.5 M Li Sulfate	none	15 % w/v PEG 8000
51.	2.0 Sodium chloride	none	10 % w/v PEG 6000
52.	0.01 M Cetyl trimethylammoniumbromide	none	0.5 M Na chloride 0.01 M Magnesium chloride hexahydrate
53.	none	none	25 % v/v Ethylene glycol
54.	none	none	35 % v/v Dioxane
55.	2.0 M Ammonium sulfate	none	5 % v/v isopropanol
56.	none	none	1.0 M Imidazole pH 6.5
57.	none	none	10 % w/v PEG 1000 10 % w/v PEG 8000
58.	1.5 M Na chloride	none	10 % v/v Ethanol
59.	none	0.1 M Na acetate pH 4.6	2.0 M Na chloride
60.	0.2 M Na chloride	0.1 M Na acetate pH 4.6	30 % v/v MPD
61.	0.01 M Cobalt chloride hexahydrate	0.1 M Na acetate pH 4.6	1.0 M 1,6 Hexanediol
62.	0.1 M Ca chloride	0.1 M Na acetate pH 4.6	30 % v/v PEG 400
63.	0.2 M NH_4 sulfate	0.1 M Na acetate pH 4.6	30 % w/v PEG 400
64.	0.2 M NH_4 suflfate	0.1 M Na citrate pH 5.6	2.0 M NH_4sulfate
65.	0.5 M NH_4 sulfate	0.1 M Na citrate pH 5.6	1.0 M Li sulfate
66.	0.5 M Na chloride	0.1 M Na citrate pH 5.6	4 %w/v Polyethyleneimine
67.	none	0.1 M Na citrate pH 5.6	35 % v/v tert-butanol
68.	0.01 M Ferric chloride hexahydrate	0.1 M Na citrate pH 5.6	10 % v/v Jerramine M-600
69.	none	0.1 M Na citrate pH 5.6	2.5 M 1,6 Hexanediol
70.	none	0.1 M MES pH 6.5	1.6 M Magnesium sulfate heptahydrate

Table 3.1: —*continued.*

71.	0.1 M Na phosphate monobasic 0.1 M K phosphate monobasic	none	none
72.	none	0.1 M MES pH 6.5	12 % w/v PEG 20,000
73.	1.6 M NH_4 sulfate	0.1 M MES pH 6.5	10 % v/v Dioxane
74.	0.05 M Cesium chloride	0.1 M MES pH 6.5	30 % v/v Jeffamine M-600
75.	0.01 M Cobalt chloride hexahydrate	0.1 M MES pH 6.5	1.8 M NH_4 suflate
76.	0.2 M NH_4 suflate	0.1 M MES pH 6.5	30 % w/v PEG monomethylether
77.	0.01 M Zinc sulfate heptahydrate	0.1 M MES pH 6.5 0.1 M MES pH 6.5	25 % v/v PEG monomethylether 550
78.	none	none	1.6 M Na citrate pH 6.5
79.	0.5 M NH_4 sulfate	0.1 M Hepes pH 7.5	30 % v/v MPD
80.	none	0.1 M Hepes pH 7.5	10 % wv/ PEG 6000 5 % v/v MPD
81.	none	0.1 M Hepes pH 7.5	20 % v/v Jeffamine M-600
82.	0.1 M Na chloride	0.1 M Hepes pH 7.5	1.6 M NH_4 sulfate
83.	none	0.1 M Hepes pH 7.5	2.0 M NH_4 formate
84.	0.05 M cadmium sulfate octahydrate	0.1 M Hepes pH 7.5	1.0 M Na acetate
85.	none	0.1 M Hepes pH 7.5	70 % v/v MPD
86.	none	0.1 M Hepes pH 7.5	4.3 M Na chloride
87.	none	0.1 M Hepes pH 7.5	10 % w/v PEG 8000 8 % v/v ethylene glycol
88.	none	0.1 M Hepes pH 7.5	20 % w/v PEG 10,000
89.	0.2 M Magnesium chloride hexahydrate	0.1 M Tris pH 8.5	20 % w/v PEG 3.4 M 1,6 Hexanediol
90.	0.1 M Ca chloride dihytdrate	0.1 M Tris pH 8.5	25 % v/v tert-butanol
91.	0.01 M Nickel chloride hexahydrate	0.1 M Tris pH 8.5	1.0 M Li sulfate
92.	1.5 M NH_4 sulfate	0.1 M Tris pH 8.5	12 % v/v glycerol
93.	0.2 M NH_4 phospharte monobasic	0.1 M Tris pH 8.5	50 % v/v MPD
94.	none	0.1 M Tris pH 8.5	20 % v/v Ethanol
95.	0.01 M Ni chloride hexahydrate	0.1 M Tris pH 8.5	20 % w/v PEG monomethylether 2000
96.	0.1 M Na chloride	0.1 M Bicine pH 9.0	30 % w/v PEG monomethylether 550
97.	none	0.1 M Bicine pH 9.0	2.0 M Magnesium chlroride hexahydrate
98.	2 % v/v Dioxane	0.1 M Bicine pH 9.0	10 % w/v PEG 20,000

The reservoir conditions used are listed in Table 3.1. Visible but thin crystals were obtained from no.17, no.28, no.31 and no.29 under the conditions listed in Table 3.1. After some trials to enlarge the crystals obtained, taking account of the successful conditions in Table 3.1, prismatic crystals were harvested, as shown in Figure 3.1 which differed in shape from those of the native crystals. The crystallographic parameters of the complex is

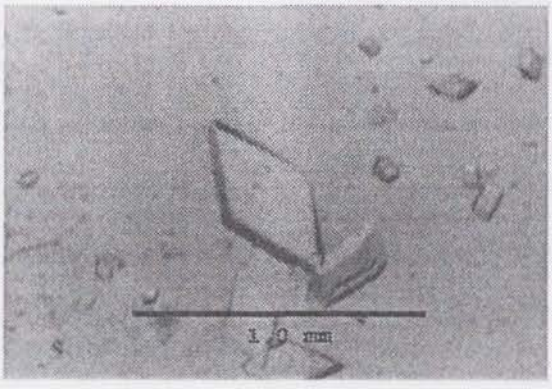


Figure 3.1: Crystals prepared by the co-crystallization.

summarized in The data from a co-crystallized complex was collected under the condition shown in Table 3.2.

Soaking in an inhibitor solution

Native crystals were soaked in ligand solutions to prepare complexes. An inhibitor solution was prepared as follows;

1. 7.37 mg of the inhibitor was dissolved in 200 μl of acetone in a tube (1.5 ml volume).
2. 200 μl of the reservoir inhibitor solution was added to the inhibitor solution, then the mixed solution was sonicated for 30 min.
3. 200 μl of acetone and 400 μl of the reservoir solution were added to the sonicated solution, which was sonicated further for 30 min.
4. A native crystal was soaked in a siliconized capillary, half of which was filled with the solution (10 mM inhibitor, acetone + the reservoir solution (3:2 v/v) solution), for 5 days at 25 °C in a procedure similar to that described in Figure 2.11, page 23.

[1]Because the inhibitor is slightly soluble in water, DMSO was added to the solution to increase the solubility.

Soaking in synthetic peptide solutions

Three synthetic peptides were kindly donated by Dr. K. Morihara. They are;

Z-Ala-Gly-Gly-Leu (Mw = 492)
Z-Phe-Gly-Ala-Ala (Mw = 523)
Z-Phe-Lys-Ala-Ala (Mw = 612)

However, both Z-Phe-Gly-Ala-Ala and Z-Phe-Lys-Ala-Ala were slightly soluble in the reservoir solution, so Z-Ala-Gly-Gly-Leu was used for further experiments as follows;

1. 4.92 mg of the Z-Ala-Gly-Gly-Leu was dissolved in 1 ml of the reservoir solution (50 mM acetate buffer (pH 5.6), 6 % (w/v) polyethylene glycol 6000 and 1 mM NaN_3), to prepare a 10.0 mM Z-Ala-Gly-Gly-Leu solution.

2. A native crystal was soaked in the solution, in a similar manner to that described above, for 3 days at 25 °C.

Besides them, another synthetic peptide;

Ac-Tyr-Val-Gly (Mw = 333, Peptide Institute Inc., Osaka, Japan)

was purchased and tried. This peptide has the virtu of being soluble in the reservoir solution with acetone. A native crystal was soaked in the peptide solution prepared as follows;

1. 3.3 mg of the specimen was dissolved in 250 μl of the reservoir solution (50 mM acetate buffer (pH 5.6), 6 % (w/v) polyethylene glycol 6000 and 1 mM NaN_3).

2. 500 μl of acetone was added to the solution to increase the solubility of the peptide, and the solution was stirred for 30 min.

3. A native crystal was soaked in a siliconized capillary, half of which was filled with the solution at 25 °C for 3 days, in a similar way to that in Figure 2.11, page 23.

3.3 Data collection of complexes

Diffraction data the complexes were collected using an R-AXIS IIC (Rigaku Denki Co., Ltd., Tokyo, Japan). Data collection conditions for the complexes prepared by the soaking and co-crystallization methods are shown in Table 3.2, and soaked ones in Table 3.3—3.5, respectively. All data sets were

Table 3.2: Data collection condition for an inhibitor complex prepared by co-crystallizaion

Ligand (co-crystallized)	Bz-Phe-Arg-(NH)-(C_6H_4)-*m*-Bzl
Crystal size	0.5×0.4×0.3 mm
Wavelength	1.5418 Å (Cu-Kα) Å
Measurement apparatus	R-AXIS IIC
Generator	RU-300
Focus	0.3×3.0 mm, double mirror focusing, fine
Filter	Ni, 0.00015 mm
X-ray power	40 kV, 100 mA
Rotation axis	$\cong b^*$
Oscillation angle	2.0 deg.
Number of flames	47
Crystal-to-detector distance	170.52 mm
Detector two-theta angle	12.99 deg.
Exposure time per flame	25 min.
Resolution	2.2 Å
Measurement temperature	20.0 °C

Table 3.3: Data collection conditions for an inhibitor complex prepared by the soaking. A native crystal was soaked in a 10 mM Bz-Phe-Arg-(NH)--(C_6H_4)-*m*-Bzl solution for 5 days on 25°C.

Ligand (soaked)	Bz-Phe-Arg-(NH)-(C_6H_4)-*m*-Bzl
Crystal size	1.0×0.6×0.4 mm
Wavelength	1.5418 Å (Cu-kα) Å
Measurement apparatus	R-AXIS IIC
Generator	RU-300
Focus	0.3×3.0 mm, double mirror focusing, fine
Filter	Ni, 0.00015 mm
X-ray power	40 kV, 100 mA
Rotation axis	$\cong b^*$
Oscillation angle	2.0 deg.
Number of flames	52
Crystal to detector distance	169.54 mm
Detector two-theta angle	13.93 deg.
Exposure time per flame	25 min.
Resolution	2.1 Å
Measurement temperature	20.0 °C

Table 3.4: Data collection condition for Z-Ala-Gly-Gly-Leu complex prepared by soaking. A native crystal was soaked in a 10 mM Z-Ala-Gly-Gly-Leu solution for 3 days at 25 °C.

Ligand (soaked)	Z-Ala-Gly-Gly-Leu
Crystal size	0.5×0.4×0.3 mm
Wavelength	1.5418 (Cu-kα) Å
Measurement apparatus	R-AXIS IIC
Generator	RU-300
Focus	0.3×3.0 mm, double mirror focusing, fine
Filter	Ni, 0.00015 mm
X-ray power	40 kV, 100 mA
Rotation axis	$\cong b^*$
Oscillation angle	2.0 deg.
Number of flames	52
Crystal to detector distance	171.50 mm
Detector two-theta angle	14.06 deg.
Exposure time per flame (mim.)	25 min.
Resolution	2.1 Å
Measurement temperature	20.0 °C

Table 3.5: Data collection condition for Ac-Try-Val-Gly complex. A native crystal was soaked into a 10 mM Ac-Try-Val-Gly solution for 3 days at 25 °C.

Ligand (soaked)	Ac-Try-Val-Gly
Crystal size	1.0×0.6×0.2 mm
Wavelength	1.5418 (Cu-Kα)
Measurement apparatus	R-AXIS IIC
Generator	
Focus	0.3×3.0, double mirror focusing, fine
X-ray power	40 kV, 100 mA
Rotation axis	$\cong b^*$
Oscillation angle	2.0
Number of flames	47
Crystal to detector distance	170.52 mm
Detector two-theta angle	12.99 deg.
Exposure time per flame	25 min.
Resolution	2.2 Å
Measurement temperature	20.0 °C

processed using a profile fitting method [137] to generate structure factor amplitudes with a program R-AXIS IIC Data Processing Software installed on a workstation Iris Indigo2 (Silicon Graphics Inc., California, U. S. A.). Data collection statistics for the complexes are shown in Table 3.7.

Molecular replacement

The molecular replacement method [138] is one of the ways to phase structure factor amplitude |F|(h k l)|s. The method assumes that proteins, which are homologous in their amino acid sequences, are similar in the tertiary structures, even if their crystal systems differ each other. A structurally established molecule can be placed in the target unit cell by determining the proper orientation and position. First, the proper rotation function must be calculated by rotation search. The next step is a translation search to superimpose the known molecule on the unknown. The molecular replacement method was performed with the program X-PLOR version 3.1 based on Patterson correlation (PC) refinement [139], using the refined model of the alkaline protease at 2.0 Å resolution as an initial searching model. The procedures were as follows;

1. Water molecules and calcium ions were removed from the initial search model.

2. The initial rotation functions were calculated with the Eulerian angles (θ_1 θ_2 θ_3), d using a direct rotation search protocol in the X-PLOR package.

3. The most probable 160 sets of Eulerian angles were selected according to their Patterson correlation coefficient values.

4. The translation function was calculated searching for the asymmetric unit of the target unit cell ($0 \leq x, y, z \leq 0.5$) divided into multiple grids in dimensions of 1.0×1.0×1.0 Å.

5. The initial searching model was superimposed well on a molecule in the target unit cell according to the result from the rotation and translation functions.

6. The generated model was refined using the slow-cooling protocol [84, 131] in the X-PLOR package, adding water molecules to the model and gradually increasing the resolution to 2.5 Å.

Prior to the structure refinements of the complexes prepared by the soaking method, a structure model of each ligand was constructed referring to a parameter handbook [140]. The structure of each complex was refined by the

Table 3.6: Crystallization conditions and crystallographic parameters of complexes prepared by co-crystallization. The parameters were refined with an R-AXIS IIC. The crystal contains two molecules in the asymmetric unit.

Protein solution	2 % (w/v) protein solution of 5 mM $CaCl_2$/1 mM NaN_3 + 1 % dimethylsulphoxid (DMSO) + 1 mM synthetic inhibitor
Incubation	37°C, 12 hours
Method	sitting drop vapor diffusion
Time (month)	1–2
Crystal size (mm)	0.5×0.4×0.3
Space group	$P2_12_12_1$
a (Å)	81.8
b (Å)	164.2
c (Å)	80.1
Z	8
V_m (Å^3)	1075,830.5
V_M (Å)	2.69
V_{solv}	0.46

Table 3.7: Data statistics of complexes.

	Crystals			
Ligand	Inhibitor		Z-Ala-Gly-Gly-Leu	Ac -Tyr-Val-Gly
Preparation method	*co-cry.	soaking	soaking	soaking
Resolution (Å)	15–2.1	15–2.1	15–2.1	15–2.1
Total observations ($I > \sigma(I)$)	45,719	34,437	35,437	33,935
Unique reflections	22,926	23,458	24,017	26,684
Completeness of data (%)	96.8	81.3	80.1	81.8
[a] R_{merge} (%)	10.03	7.45	6.88	6.40
$R_{\mathrm{reliability}}$ (%)	9.61	7.95	7.33	7.17

[a] $R_{\mathrm{merge}}(\%) = 100\Sigma_i|\langle I_i\rangle - I_i|/|\Sigma i I_i$;
$\langle \mathrm{I_i}\rangle$ is the average of $\mathrm{I_i}$ over all observed symmetry equivalents.
*co-crystallization.

slow-cooling protocol [84, 131] at 2.3 Å resolution. The refinement procedures were as follows;

1. The model was refined by one cycle of the slow-cooling protocol. The annealing temperature during the refinement was reduced from 800 K to 300 K.

2. A $|\mathbf{F}_{\mathrm{obs}}| - |\mathbf{F}_{\mathrm{calc}}|$ (Difference Fourier) map of the complexe was calculated, and each ligand model was superimposed on the electron density map, removing the water molecules around the placed ligands.

3. In regions where large conformational changes occurred, the model was manually rebuilt to fit the corresponding $|\mathbf{F}_{\mathrm{obs}}| - |\mathbf{F}_{\mathrm{calc}}|$ maps.

4. The second cycle of slow-cooling was performed with the modified model by reducing the temperature from 800 K to 300 K.

PC-refinement and translation search

The direct rotation search followed by the *PC*-refinement [139] generated six peaks (> 0.14 in Patterson correlation coefficient) as shown in Figure 3.2. Each set of the Eulerian angles derived from the peaks of the rotation function was used for setting the initial orientation of the search model to perform the following translation searches, which were carried out for each

of two molecules in the asymmetric unit.

The resulting crystal packings were examined to see whether the arrange-

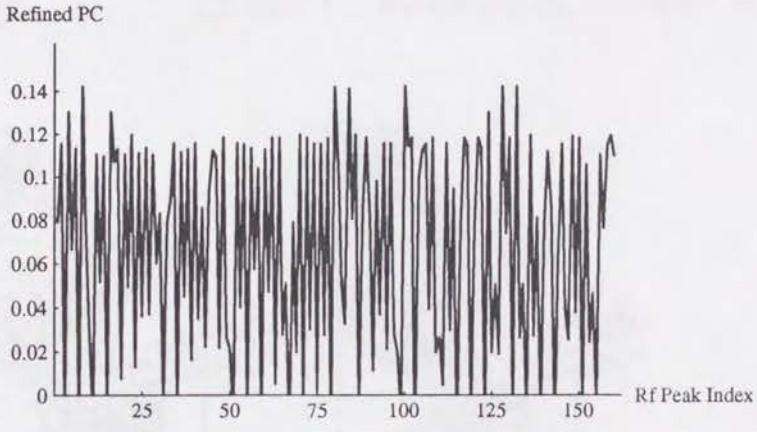


Figure 3.2: PC-refinement. Peaks come from Patterson correlation (PC) refinement. Six peaks showing values larger than 0.14 in patterson correlation coefficient were found. The vertical axis indicates values of PC, and the horizontal, solution indices of the preceded rotation search.

Table 3.8: Results from the rotation and translation functions. The searching operations were carried out independently for each of two molecules in the asymmetric unit.

	molecule A	molecule B
Rotation function (°) $(\theta_1, \theta_2, \theta_3)$	(98.714 56.485 137.58)	(93.125 41.904 141.24)
Translation function (Å)	(7.775 0.000 12.411)	(35.805 72.850 22.520)
Patterson correlation	0.1976	0.1974

ment of the molecules in the unit cell was appropriate, and the crystallographic statistics calculated from the packings matched the actual ones. Finally, one set of the rotational and translational parameters could be selected among the candidates. This set of functions is listed in Table 3.8. The structure model of two anhydrous molecules, reasonably located in the unit cell, was refined with a simulated annealing protocol in the program package X-PLOR (version 3.1), and water molecules were added to the model.

The current crystallographic R-value of the structure, which consists of two enzyme molecules and 195 water molecules, has been reduced to 21.2 % for the data between 8.0–2.8 Å resolution. The refinement procedure is now in progress, by gradually extending the resolution. The mode of crystal packing of the complex is demonstrated in Figure 3.3.

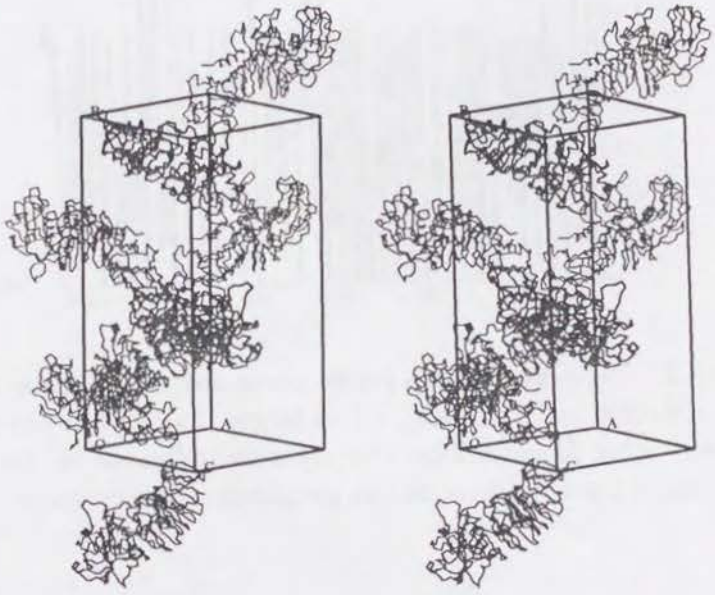


Figure 3.3: Crystal packing of a complex obtained by the co-crystallization. Each asymmetric unit contains two molecules. The crystal belongs to space group, $P2_12_12_1$, with unitcell dimensions of a = 81.8 Å, b = 164.2 Å, and c = 80.1 Å.

3.3.1 Crystal structures of complexes prepared by the soaking method

Complexes prepared by the soaking method turned out to be more suitable for complex analysis than those prepared by the co-crystallization because the isomorphism of those prepared by the soaking method was completely maintained after ligand binding. Structure refinement of the three complexes carried out using a few cycles of the simulated annealing technique, complemented by manual model re-building and interpretation of water molecules. The final model statistics of the complexes are shown in Table 3.9. The difference Fourier maps and schematic representations of ligands are depicted in Figure 3.4—3.6. The refined coordinates of ligands are on Table 3.10—3.12,

respectively. Ligand–protein distances are summarized on Table 3.13. To

Table 3.9: Final model statistics of complexes prepared by soaking method.

	Crystals		
	Inhibitor	Z-Ala-Gly-Gly-Leu	Ac-Tyr-Val-Gly
Resolution (Å)	8.0–2.3	8.0–2.3	8.0–2.2
No. of reflection	21,118	21,270	24,834
Completeness of data (%)	67.8	68.3	69.3
R-value (%)	18.0	17.7	18.8
rmsd in			
bond length (Å)	0.012	0.026	0.026
angle (°)	2.896	1.790	1.866
No. of water molecules	289	300	289

confirm the ligand binding to the active site of the alkaline protease, difference Fourier maps were calculated with the refined models. The difference Fourier maps revealed clear electron densities for the ligands as shown in Figure 3.4–3.6.

3.4 Discussion

As shown in Figure 2.29, page 60, the portion corresponding to the flexible loop region of the alkaline protease, Tyr190–Glu195, are almost completely conserved for all members of the metzincin superfamily, which is consistent with the observation that the flexible loop region of the alkaline protease plays a significant role in ligand stabilization as mentioned above. Tyrosine residues which correspond to Tyr216 of the alkaline protease are present in the conserved Met-turn of the superfamily . The tyrosines probably serve as the fifth ligand to the active site zinc, and their conformation should change on ligand binding in the same way similar to that observed in the alkaline protease. In conclusion, the ligand stabilization mechanism the the 'flipping' of Tyr216 and its hydrogen bonds to Asn191 in the flexible loop, may be a common mechanism among the serralysins. The so-called 'S_1' pocket' of the serralysins are not hydrophobic, which is in contrast with those of the thermolysins and carboxypeptidases [141, 43]. In the alkaline protease, the S_1' pocket is formed by Thr173, Arg209, Tyr216 and six water molecules,

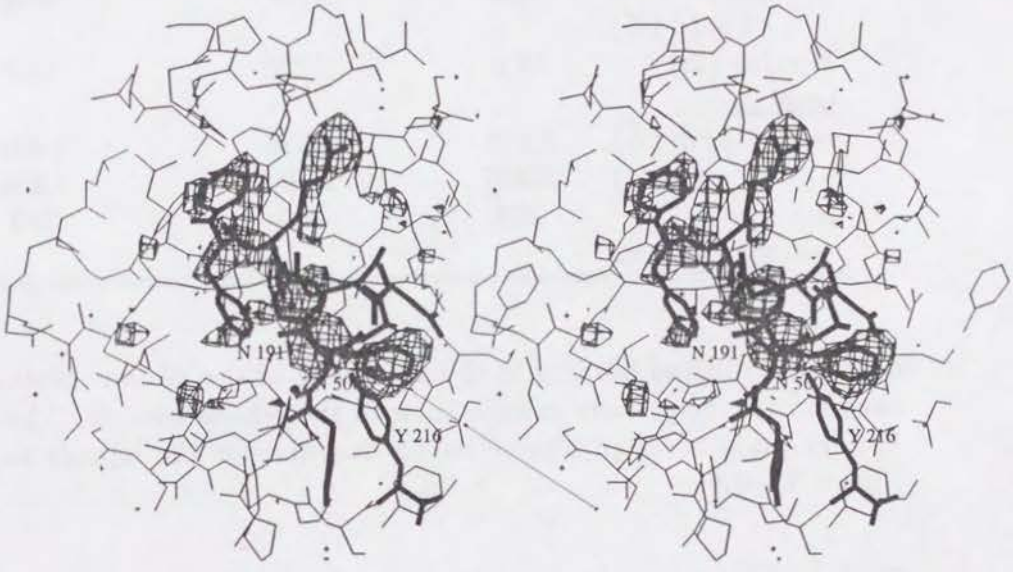


Figure 3.4: A stereo drawing of the final difference Fourier map around the inhibitor. The final refined skeletal model was superimposed on the electron density, contoured at 2.8 σ level. A flexible loop region is depicted in a bold line.

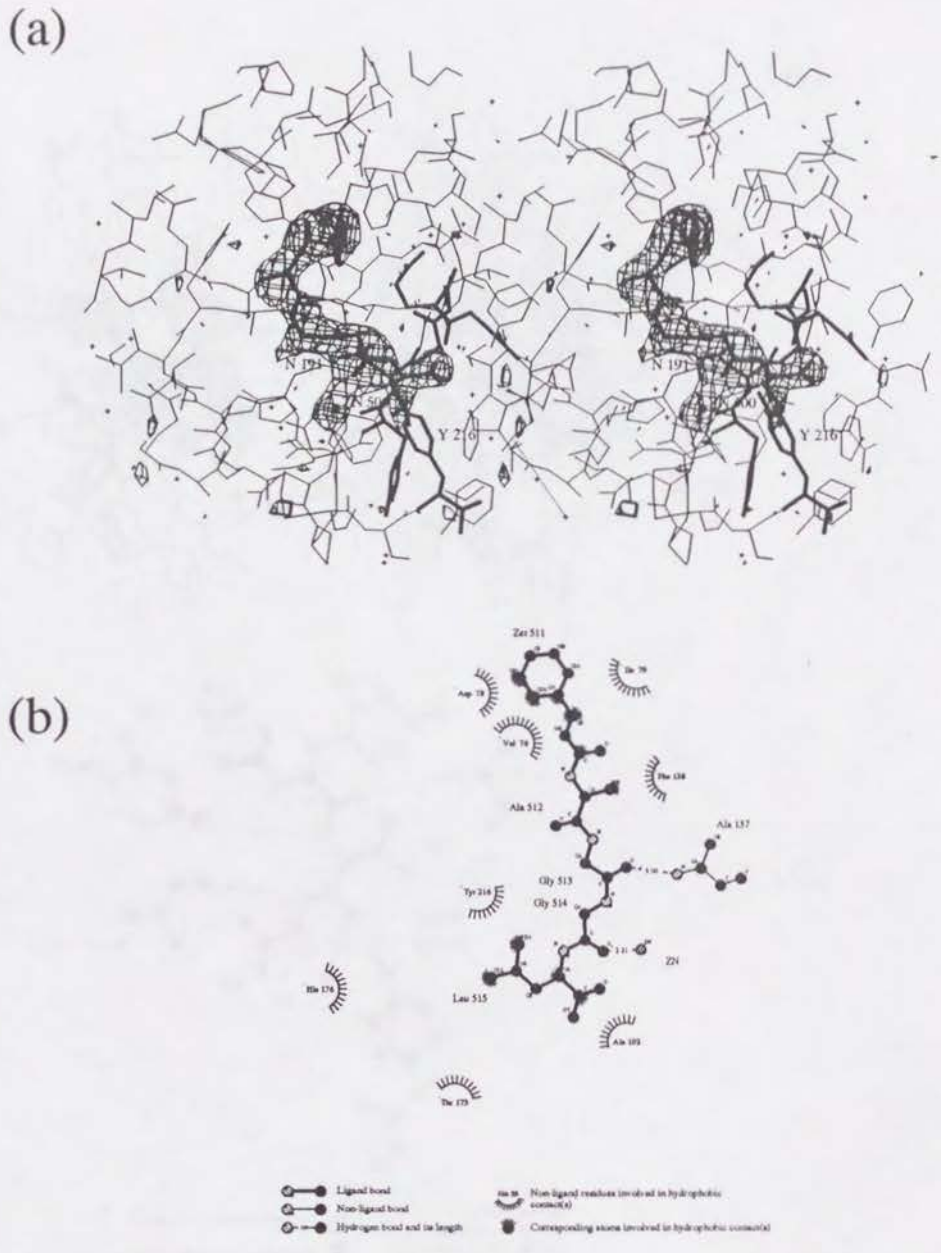
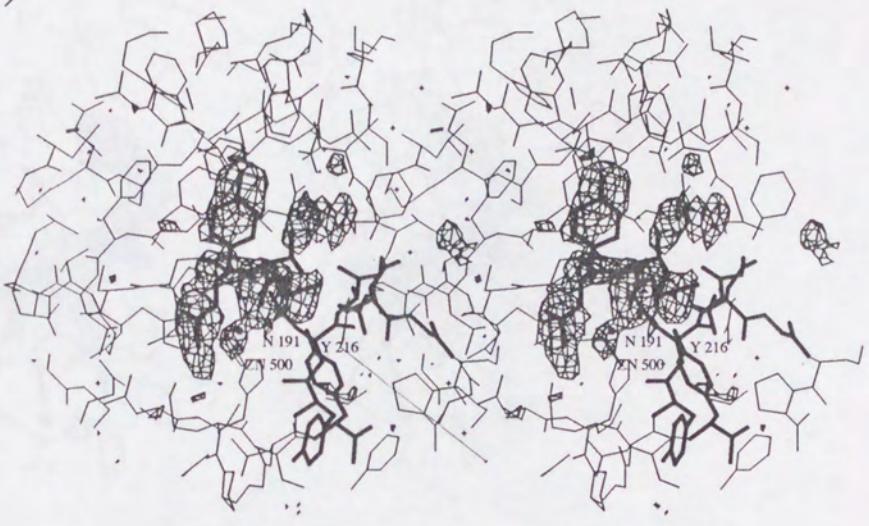


Figure 3.5: (a) A stereo drawing of the final difference Fourier map around the Z-Ala-Gly-Gly-Leu peptide ligand. The final refined skeletal model was superimposed on the electron density, contoured at 3.0 σ level. (b) a schematic representation of ligand–protein interactions.

(a)



(b)

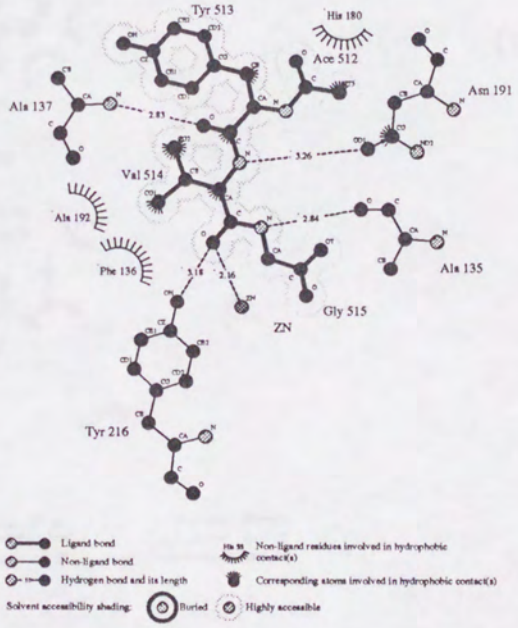


Figure 3.6: (a) A stereo drawing of the final difference Fourier map around the Ac-Tyr-Val-Gly peptide ligand, contoured at 3.0 σ level. (b) a schematic representation of ligand–protein interactions.

Table 3.10: Refined inhibitor coordinates. B; temperature factor, BZL; benzyl, MA; mercaptoamilide, BZ; benzene. *X*, *Y*, *Z* are in angstroms in an orthogonal coordinates (Å). Occupancies are set to unity.

Atom	Residue	*X* (Å)	*Y* (Å)	*Z* (Å)	B (Å^2)
1CB	BZL	55.822	82.281	24.953	54.97
1CG	BZL	56.614	81.496	25.967	53.53
1CD1	BZL	56.480	81.805	27.317	51.99
1CD2	BZL	57.480	80.472	25.600	53.10
1CE1	BZL	57.188	81.098	28.278	51.78
1CE2	BZL	58.196	79.748	26.563	50.93
1CZ	BZL	58.052	80.064	27.898	51.86
2SD	MA	56.103	81.933	23.149	51.57
2CG	MA	55.294	83.304	22.188	59.35
2CD1	MA	55.111	83.182	20.805	59.63
2CD2	MA	54.867	84.419	22.892	58.97
2CE1	MA	54.428	84.257	20.120	59.41
2CE2	MA	54.208	85.438	22.235	58.54
2CZ	MA	53.972	85.379	20.857	59.86
2N	MA	53.268	86.468	20.218	61.21
N	ARG	50.238	87.381	18.572	58.35
CA	ARG	51.343	87.717	19.472	58.72
CB	ARG	50.863	88.641	20.617	57.15
CG	ARG	50.311	90.004	20.171	56.87
CD	ARG	51.333	91.133	20.317	56.17
NE	ARG	51.386	91.664	21.685	57.23
CZ	ARG	51.731	92.917	22.006	56.73
NH1	ARG	52.064	93.791	21.064	57.42
NH2	ARG	51.766	93.296	23.280	58.06
C	ARG	51.930	86.440	20.085	59.70
O	ARG	51.284	85.380	19.939	57.28
3N	PHE	48.470	88.353	15.570	60.71
3CA	PHE	49.032	87.334	16.461	59.72
3CB	PHE	49.555	86.156	15.636	59.13
3CG	PHE	49.982	84.978	16.463	60.27
3CD1	PHE	49.042	84.021	16.869	59.89
3CD2	PHE	51.336	84.788	16.793	59.85
3CE1	PHE	49.439	82.876	17.591	60.29
3CE2	PHE	51.753	83.651	17.514	60.20
3CZ	PHE	50.798	82.689	17.911	60.46
3C	PHE	50.161	87.891	17.343	59.52
3O	PHE	51.022	88.660	16.871	59.61
4CG	BZ	47.192	88.997	15.911	60.36
4CD1	BZ	46.866	89.283	17.218	59.69
4CD2	BZ	46.273	89.325	14.869	59.39
4CE1	BZ	45.664	89.932	17.539	58.60
4CE2	BZ	45.070	89.980	15.174	58.58
4CZ	BZ	44.762	90.281	16.515	58.97

Table 3.11: Refined ZAGGL ligand coordinates. B; temperature factor, Z; benzyloxycarbonyl. *X*, *Y*, and Z are in the orthogonal coordinates (Å). Occupancies are set to unity.

Atom	Residue	*X* (Å)	*Y* (Å)	*Z* (Å)	B (Å^2)
CB	Z	54.206	89.870	19.442	51.47
CG	Z	52.808	89.804	18.917	52.35
CD1	Z	52.143	88.575	18.833	52.92
CD2	Z	52.150	90.962	18.507	52.56
CE1	Z	50.834	88.505	18.344	52.64
CE2	Z	50.850	90.913	18.015	51.62
CZ	Z	50.187	89.683	17.931	51.98
OE	Z	54.972	88.900	19.432	51.13
C	Z	56.088	88.932	18.668	51.89
O	Z	56.986	89.734	18.928	50.81
N	ALA	56.350	87.633	18.202	49.12
CA	ALA	55.404	86.890	17.374	50.46
CB	ALA	56.109	86.377	16.148	49.90
C	ALA	54.832	85.720	18.177	50.90
O	ALA	53.818	85.858	18.887	51.92
1N	GLY	55.525	84.504	18.041	52.07
1CA	GLY	55.114	83.324	18.775	50.58
1C	GLY	55.874	83.384	20.091	50.74
1O	GLY	57.108	83.264	20.089	49.35
2N	GLY	55.114	83.594	21.264	51.54
2CA	GLY	55.741	83.698	22.578	53.35
2C	GLY	55.884	82.437	23.406	53.83
2O	GLY	56.124	81.360	22.864	52.20
N	LEU	55.785	82.614	24.804	56.50
CA	LEU	55.860	81.532	25.807	59.58
CB	LEU	57.253	80.872	25.848	57.29
CG	LEU	57.343	79.775	26.927	56.23
CD1	LEU	56.168	78.803	26.759	53.61
CD2	LEU	58.650	79.034	26.873	54.79
C	LEU	55.474	82.010	27.238	61.87
O	LEU	54.280	81.883	27.609	63.11
OT	LEU	56.365	82.477	27.994	63.32

Table 3.12: Refined Ac-Y-V-G ligand coordinates. B; temperature factor, AC; acetyl. Occupancies are set to unity.

Atom	Residue	X (Å)	Y (Å)	Z (Å)	B (Å^2)
C	AC	54.448	81.572	17.358	52.88
O	AC	54.464	82.160	16.268	54.31
CT3	AC	53.775	80.195	17.475	51.45
N	TYR	54.589	82.244	18.586	51.52
CA	TYR	55.309	83.523	18.701	50.08
CB	TYR	54.316	84.679	18.518	49.36
CG	TYR	54.889	86.066	18.692	50.50
CD1	TYR	55.760	86.616	17.740	49.27
CE1	TYR	56.261	87.923	17.893	49.75
CD2	TYR	54.534	86.853	19.802	50.25
CE2	TYR	55.028	88.160	19.960	49.54
CZ	TYR	55.887	88.682	19.005	48.82
OH	TYR	56.365	89.958	19.160	49.67
C	TYR	56.044	83.606	20.054	48.76
O	TYR	57.259	83.373	20.108	45.72
N	VAL	55.259	83.904	21.204	48.85
CA	VAL	55.827	83.978	22.581	49.06
CB	VAL	54.994	84.961	23.517	48.11
CG1	VAL	55.360	86.468	23.285	45.98
CG2	VAL	53.517	84.746	23.290	46.76
C	VAL	55.938	82.587	23.279	48.17
O	VAL	55.532	81.582	22.717	44.57
N	GLY	56.497	82.525	24.566	50.68
CA	GLY	56.614	81.242	25.269	54.04
C	GLY	56.959	81.232	26.770	55.58
O	GLY	56.282	81.946	27.562	55.56
OT	GLY	57.899	80.490	27.167	55.60

which implies that the inside of the pocket is rather hydrophilic. Moreover, the pocket is larger than those of thermolysins and collagenases.

To the author's knowledge, there have been no complexes between zinc metalloproteases and inert peptide substrates. except for the present structures of complexes. The present X-ray structures of the complexes are landmarks from a point of view that the peptide ligands bind to the enzyme without hydrolysis. The peptide-enzyme complexes mimic the interactions between the natural substrates and the enzyme, providing information more useful to elucidate the catalytic mechanism of the enzyme than that obtained from an inhibitor–enzyme complex. The reason why the peptide-ligands were complexed with the enzyme without hydrolyzation is thought to be as follows.

- The pH values (pH 5.6) in preparing the complexes were rather acidic compared with the optimum pH (pH7–9) of the enzyme. In such an acidic condition, the $O\delta_{1,2}$ of the 'general base residue' Glu177 were probably protonated, decreasing the nucleophilicity of the residue. In a phosphate inhibitor–thermolysin complex, Glu143, which is presumed to be a general base residue, also appears to be protonated, even in a solution at pH 7.2 [46].

- The thermal motions of the molecules packed in a crystal were more restricted by crystal packing force than those in solution, which leads to a decrease in catalytic activity.

- The temperatures applied for preparing the complexes (25 °C) and collecting data (20 °C) were lower than the optimum (60 °C).

On ligand-binding to the enzyme, most of the water molecules present in the catalytic cavity (*c.f.* Figure 2.27 page 58) were replaced by ligands, and large conformational changes occurred mainly in the side chain of Tyr216 and the flexible loop region comprising Try190–Glu195, as shown in Figure 3.7, page 115. Similar movement of a tyrosine residue near the active site zinc is observed in Tyr248 of carboxypeptidase A on the binding of a peptide in the active site [142]. In carboxypeptidase A, Tyr248 included in a flexible loop on the molecular surface is not a zinc ligand, unlike Tyr216 in the alkaline protease; the phenolic oxygen of Tyr248 is in the vicinity of the active site, being about 7 Å from the zinc ion. The crystal structures of various carboxypeptidase A–ligand complexes show that when each ligand binds to the enzyme, Tyr248 undergoes a substrate-induced conformational change resulting in displacement of the phenolic hydroxyl group from the surface of the enzyme within the hydrogen-bonding distance of the nitrogen of the peptide bond to be cleaved. In arsanilazotyrosine-248 carboxypeptidase A, it has been shown that Tyr248 interacts with the zinc ion in order to interact with the scissile bond on ligand binding [143]. There is some controversy

over the role of Tyr248 in the mechanism of action of carboxypeptidase A. However, it has been confirmed by a mutagenesis study that the phenolic hydroxyl group of Tyr248 in carboxypeptidase A does not act as a proton donor for the amino nitrogen in the cleaved peptide bond but participates in ligand binding. In this context, the role of Tyr216 in the alkaline protease is not that of a proton donor but a residue which stabilizes ligand–enzyme complexes by forming hydrogen bonds.

The sulfur atom of the inhibitor liganded to the zinc, replacing the water molecule bound to the zinc. As can be seen in Table 3.13, the distance from the sulphur atom (2SD) to the zinc is 2.12 Å, which approximately coincides with the average observed for Zn–S distances in protein structures [144, 145]. The sulfur atom also interacts with $O\epsilon_1$-177Glu which is assumed to be a general base residue. On the other hand, the sulfur (2SD)–OH216Tyr distance is 4.10 Å, which may be slightly longer in order to interact with each other. On inhibitor binding, the conformation of Tyr216 was changed. The phenolic oxygen of the residue was moved a distance of 2.56 Å from the zinc, due to rotations of the ϕ, ψ and χ angles as shown in Table 3.14. The phenolic oxygen, in turn, made hydrogen bonds with N-192Ala and $N\delta_2$-191Asn, probably to stabilize the flexible loop region. In the native form, the ϕ-ψ angles of Try216 lie slightly outside of an α-helix region in a Ramachandran plot. However, those angles in the liganded forms fall into the α-helix region, which may imply that geometric stress caused by the coordinate bond between the OH-216Tyr and the zinc was released on ligand binding.

$N\eta_{1,2}$-Arg of the inhibitor make salt bridges with $O\delta_1$-78Asp and $O\delta_2$-78Asp stabilizing the Michaelis complex, which coincides with the result obtained previously that an arginine residue present in the peptide portion of the inhibitor increases affinity to the alkaline protease [65, 66]. These conformational changes in the Try216 and the flexible loop region were also observed in the peptide ligand complexes, as shown in Figure 3.7. The peptide ligand–enzyme complexes were further stabilized by hydrogen bonds between the $O\delta_1$-Asn191 and amide nitrogen of P_1' residues as shown in Table 3.13 and Figure 3.5 –3.6. The P_1' amide nitrogen and P_2 carbonyl oxygen of the ligands made hydrogen bonds with O-135Ala and N-Ala137 of the enzyme, respectively, probably stabilizing the ligand complexes.

On the other hand, all of the N-terminal portions of the ligands, Bz-F- in the inhibitor, Z-A- in the ZAGGL, and Ac-Y- in the AcYVG, loosely interacted with the enzyme mainly through hydrophobic forces (see Figure 3.5–3.6). Fourier maps of the complexes imply that these portions of the ligands were not so significant in determining complementarities between the ligands and the enzyme. This contrasts with the extended substrate binding mode of thermolysin, as deduced from a variety of its inhibitor–enzyme complexes [47, 42]. In the thermolysin, R_2, R_1 and R_1' interact with

the protein backbone of Trp115–Ala113 as in an antiparallel pleated sheet. In human neutrophil collagenase, an extended substrate binding mode has been also proposed, where each residue in extended P_3–P_3' portion of ligands make six hydrogen bonds with active site rims.

The so called 'S_1' pocket' of the alkaline protease comprises Thr173, Arg209, six water molecules, and the Tyr216, whose conformation changes on ligand binding as mentioned earlier. It is widely accepted that in zinc metalloendoproteases, the substrate specificities of the enzymes greatly depend on the residues forming the S_1' pockets [47, 42, 48, 141, 45, 44], because the S_1' harbors the P_1' side chains of bound substrates. The size of the pocket is difficult to estimate because its boundary is ambiguous.

On the other hand, all of the S_1' pockets of other metzincins such as adamalysin [108], astacin [40, 105, 120, 146] and collagenases [106, 136, 41, 113, 107, 145], are more sterically restricted than that of the alkaline protease, as shown in Figure 2.38. In thermolysins [47, 42, 48], the S_1' pocket is also more sterically hindered than that of the alkaline protease. These observations suggest that the S_1' pocket of the alkaline protease does not have much influence in the determination of the substrate specificity, unlike the other metzincins. This lower steric hindrance in the P_1' pocket of the alkaline protease seems to correspond to the results supporting the 'broad' specificity shown by hydrolyzing the oxidized insulin B-chain and a variety of synthetic peptides [78, 12, 13].

Affinity enhancement by Lys or Arg in the P_1 site

These observations, which substantiate the fact that the alkaline protease has a 'broad' specificity, seem to contradict the results showing that the alkaline protease specifically recognizes lysine or arginine in the P_1 site of the substrates [64, 65, 66]. To address this discrepancy, the glycine residue which corresponds to the P_1 residue of the ligand ZAGGL in the complex was replaced with lysine, and the modified complex was carefully inspected to determine whether the lysine could interact with the enzyme without steric hindrance. The result is shown in Figure 3.8. As seen in this figure, the side chain of the replaced arginine reaches the position where $N\eta_{1,2}$ of the arginine can make salt bridges with $O\epsilon_{1,2}$–Glu194 and O–Ala131. Glu194 of the enzyme, which is located in the flexible loop region, takes a relatively free conformation on ligand binding. When substrates having no basic residues (Arg or Lys) in the P_1 site bind to the enzyme, the Glu194 does not take part in ligand stabilization. However, if substrates having basic residues in the P_1 site bind to the enzyme, the side chain of Glu194 in turn enhances the ligand affinity to make salt bridges with the basic residues. In other words, although the alkaline protease can make Michaelis complexes with a variety of substrates regardless of the property of P_1 residues, as shown by

Table 3.13: Ligand-protein distances.

Ligand	Protein	Distance (Å)
Inhibitor		
2SD	ZN	2.12
2SD	$O\epsilon_1$-177Glu	3.12
2SD	$O\epsilon_2$-177Glu	3.88
2SD	OH-Tyr216	4.10
$N\eta_1$-Arg	$O\delta_1$-78Asp	2.96
$N\eta_1$-Arg	$O\delta_2$-78Asp	3.65
$N\eta_2$-Arg	$O\delta_1$-78Asp	3.74
$N\eta_2$-Arg	$O\delta_2$-78Asp	3.29
Z-A-G-G-L		
O-Gly_2	ZN	2.21
O-Gly_1	N-137Ala	3.00
O-Gly_2	$O\delta_1$-177Glu	3.28
O-Gly_2	OH-216Tyr	3.63
N-Gly	$O\delta_1$-191Asn	2.90
N-Leu	O-135Ala	3.27
Ac-Y-V-G		
O-Val	ZN	2.16
O-Tyr	N-137Ala	2.83
O-Val_2	$O\delta_1$-177Glu	4.17
O-Val	OH-216Tyr	3.18
O-Val	$O\delta_1$-191Asn	3.26
N-Gly	O-135Ala	2.84

Table 3.14: Conformational changes of Try216 on ligand binding. The ϕ ψ angles of the ligand-binding forms are present in the α-helix region of a Ramachandran plot.

	(ϕ	ψ	χ_1)	(°)
Native	(-114	2	-81.8)	
Inhibitor	(-101	-18	-88.4)	
ZAGGL	(-107	-18	-86.8)	
AcYVG	(-106	-17	-87.1)	

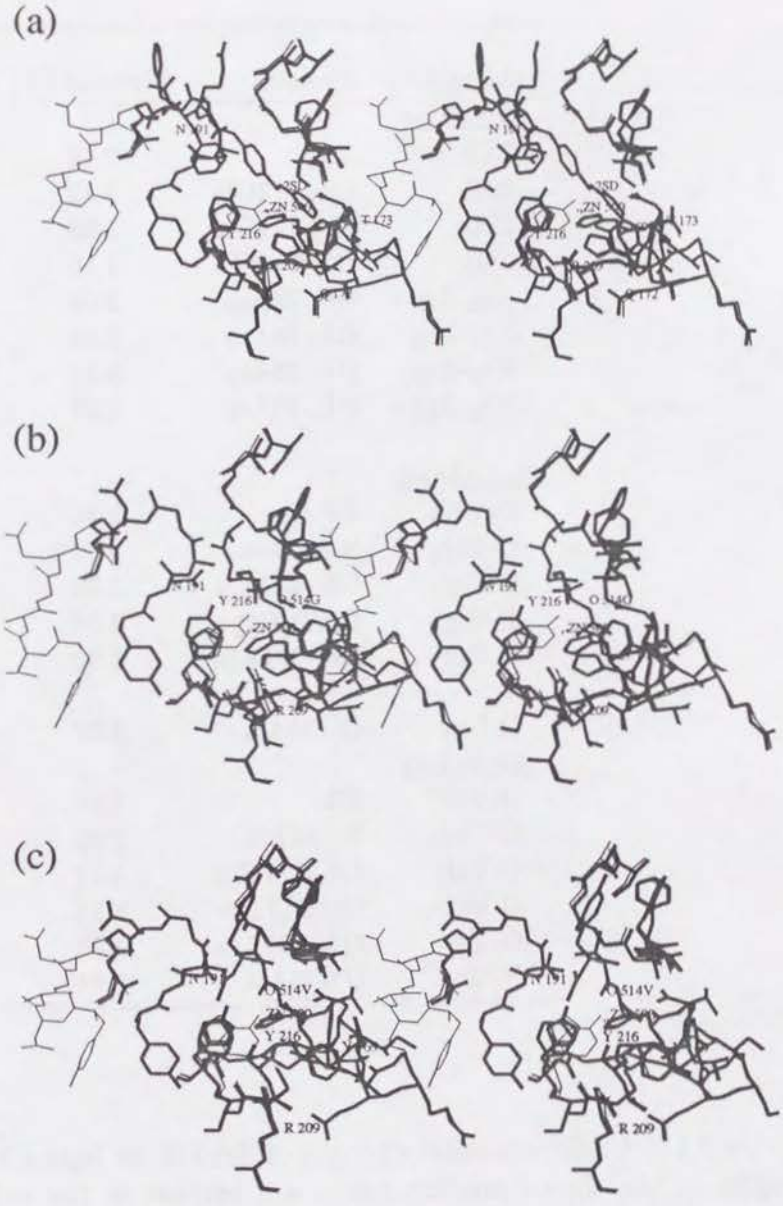


Figure 3.7: Conformational changes around the active sites on ligand binding. The skeletal models depicted in thin and bold lines indicate native and complexed forms, respectively. (a) The skeletal model of the inhibitor complex is superimposed on that of the native one. (b) ZAGGL–native. (c) AcYVG–native.

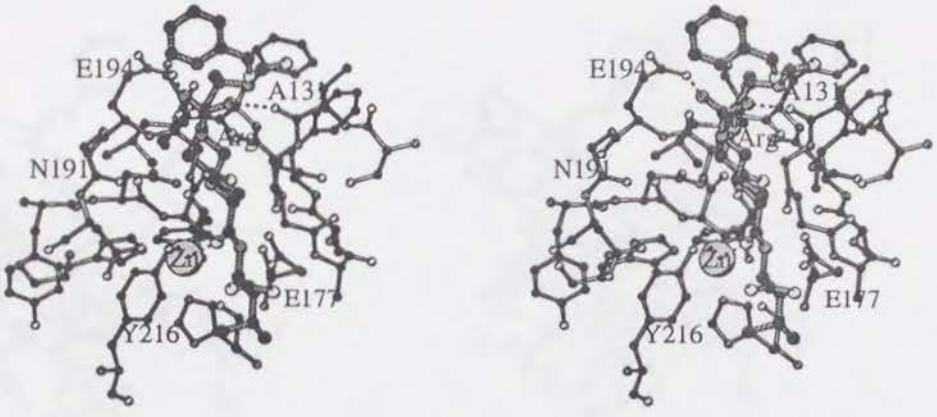


Figure 3.8: Affinity enhancement by a replaced arginine. The ligand ZAGGL, depicted in dark grey, in the complex was modified to ZAGRL. $N\eta_{1,2}$ of the arginine can make salt bridges with $O\epsilon_{1,2}$–Glu194 and O–A131.

the three kinds of complexes of the alkaline protease, the ligand affinity can be increased by the basic residue in the P_1 site. For this reason, in the series of hydrolysis experiments using fluorogenic substrates shown in Figure 2.6, page 17, it was concluded that the alkaline protease prefers arginine or lysine residues in the substrate P_1 site [64, 65, 66], resulting in a classification of the enzyme as 'COOH' type [147].

Comparison with an alkaline protease solved with a product

The crystal structure of the alkaline protease from *Pseudomonas aeruginosa* PAO1 [121] is now availabe [127]. The structure turned out to possess a tetra peptide as a product in its catalytic cleft, which may derive from impurities in the specimen. The amino acid residues of the bound peptide have been assigned in the coordinates deposited in the PDB,[2] maybe deduced from the electron density, although the literature [127] said that the ligand was a hetero tetra peptide. We can get an insight into what will occur when the two kinds of the peptide ligands, ZAGGL and AcYVG are hydrolyzed by the enzyme, based on the structure of PAO1 with the hydrolyzed product. The stereo representation of the alkaline protease from *Pseudomonas aeruginosa* PAO1 is shown in Figure 3.9 (b) together with (a) that of the ZAGGL complex. As can be seen in this figure, the tetra peptide does not take an extended conformation; instead, the N-terminal portion comprising

[2]Identity code '1KAP' in the PDB.

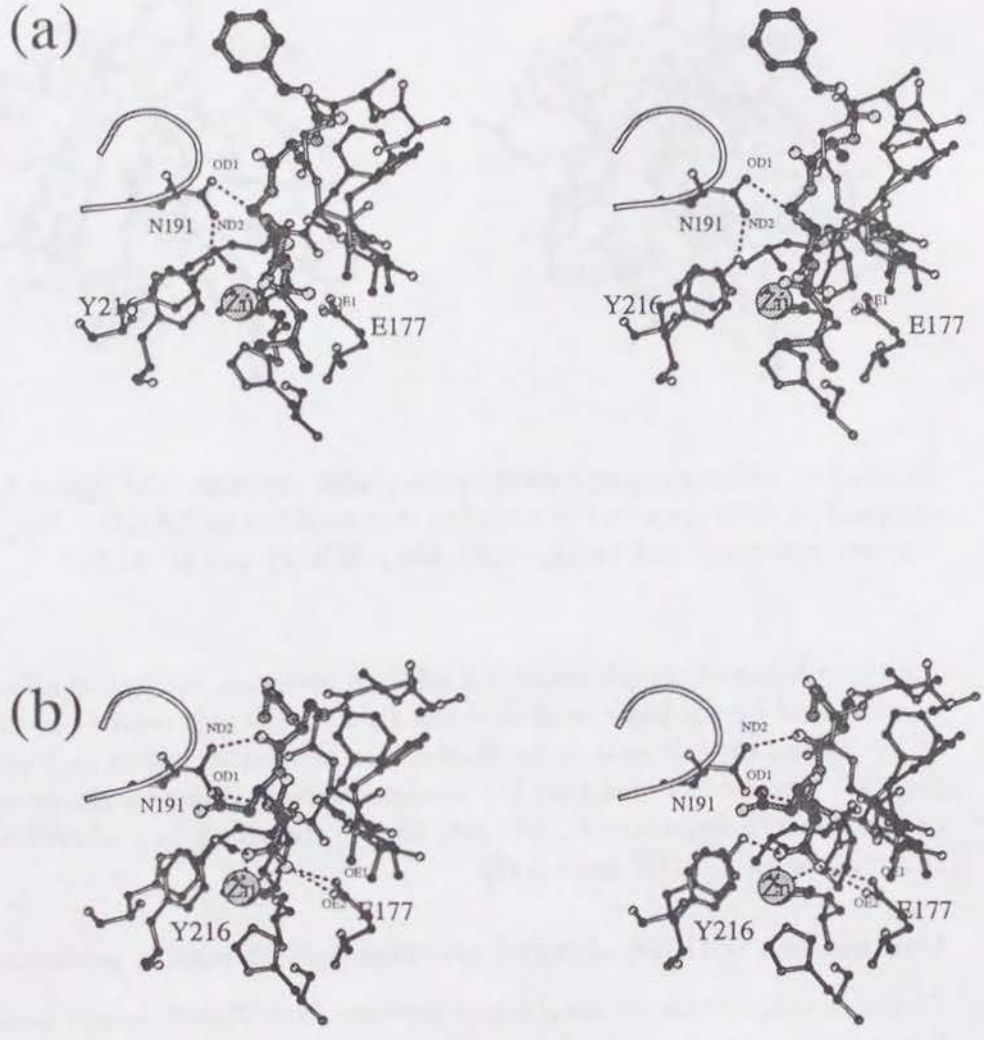


Figure 3.9: Structural changes during hydrolysis. Fixed flexible loops are represented by coils. (a) ZAGGL complex. $O\delta_1$ and $N\delta_2$ of Asn191 make hydrogen bonds with N–P_1 and OH–Tyr216, respectively, stabilizing the fixed loop conformation and the Michaelis complex. (b): GSNS complex from another form of the alkaline protease [127]. The primed portion of the peptide is thought to be cleaved by hydrolysis. As a result, the cleaved product was displaced along $\beta 6$, leading to changing conformation of the fixed flexible loop, and rotation in the side chain of Asn191 94° about the χ_2 axis. The hydrogen bond between $N\delta_2$–Asn191 and OH–Tyr216 was turned off, and, a new one between $N\delta_2$–Asn191 and O–P_3 was formed.

P_4 and P_3 residues is nearly perpendicular to $\beta 6$ without any hydrogen bonds or salt bridges with the enzyme, except that between N753I and a water molecule O630. This means that the N-terminal portion of the ligand has no significant influence on substrate recognition or binding. The situation is exactly same in the case of the ZAGGL and the AcYVG complexes of the alkaline protease, where the N-terminal portions can take relatively free conformations as shown in Figure 3.5–3.6. On the other hand, one of the two C-terminal carboxyl oxygens makes a hydrogen bond with the $O\epsilon_{1,2}$ of general base residue Glu177. The other carboxyl oxygen makes a coordination bond with the catalytic zinc, and a hydrogen bond with the phenolic oxygen of the 'flipped' Tyr216, as seen in the three kinds of ligands described in this study. The phenolic oxygen of Tyr216 also makes a hydrogen bond with the peptide nitrogen of A192 which is located in the flexible loop region. This flexible loop, however, takes a somewhat displaced conformation compared with those of the three kinds of ligand complexes in this study, in accordance with the displaced product to the 'left side' in the figure. In other words, the role of the flexible loop in stabilizing Michaelis complexes during the reaction allows considerable geometry changes in the bound ligands, which may partly explain the broad specificity of the enzyme.

In contrast, $N\delta_2$ of Asn191 cannot make a hydrogen bond with the phenolic oxygen of Tyr216, because the side chain was rotated 94° about the χ_2 axis, possibly in accordance with the displacement of the product to the 'left' side. As a result, the hydrogen bond between $N\delta_2$–Asn191 and OH–Tyr216 was turned off, the fixed flexible loop to become unstable again, and a new hydrogen bond between $N\delta_2$–Asn191 and O–P_3 residue replaces the previous one. Such an exchange of hydrogen bonds during the hydrolysis may contribute to the release of the products from the catalytic cleft.

Proposed catalytic mechanism of the alkaline protease

Taking into account the tertiary structures of the alkaline protease, and the results obtained from biophysical or biochemical experiments, the catalytic mechanism of the alkaline protease can be proposed as in Figure 3.10. In conclusion, the following catalytic mechanism of the alkaline protease is suggested. The flexible loop region of Tyr190–Gly195 may vibrate in accordance with the 'flip-flop' motion of Tyr216, to make or to break the hydrogen bond between OH-Tyr216 and $N\delta_2$-Asn191, which may contribute to increase substrate–enzyme affinity. On ligand binding, Try216, which has been the fifth ligand of the zinc, is pushed out of the S_1' pocket to make a hydrogen bond with the Asn191. The water molecule bound to the zinc ion is also pushed toward the carboxyl oxygen of Glu177 to have the enhanced nucleophilicty. The bound ligand is mainly stabilized by the coordinate bond between the P_1 carbonyl oxygen and the zinc ion, and further by O–Ala135,

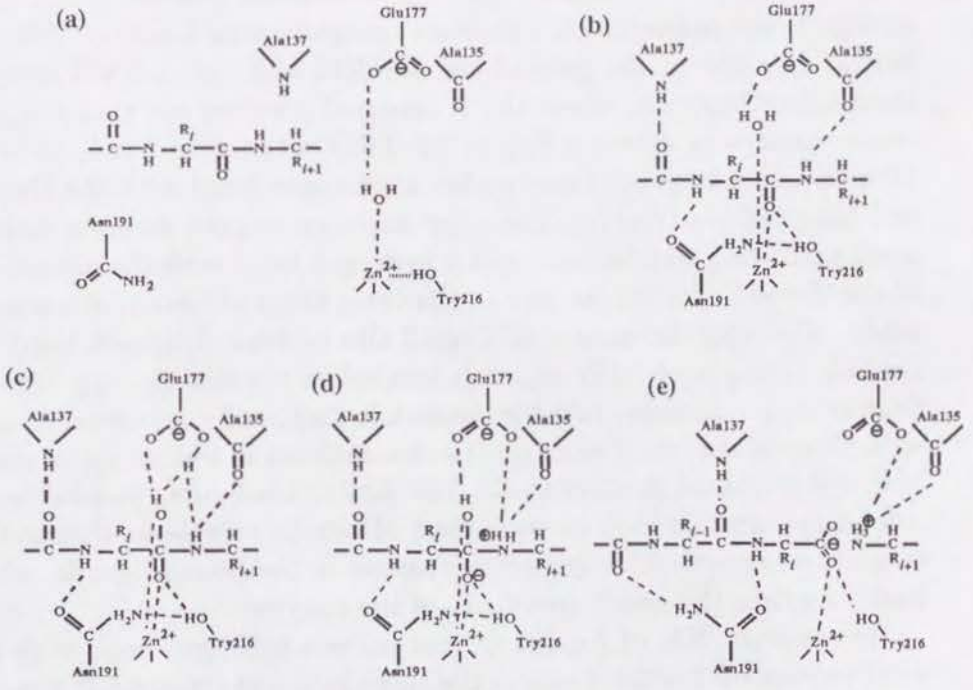


Figure 3.10: Proposed catalytic mechanism of the alkaline protease. (a) In the ligand-free form of the alkaline protease, Tyr216 and the flexible loop region of Tyr190–Gly195 which contains Asn191 and Asp194, probably join together to make a 'back and forth' or 'flip-flop' motion, mainly driven by their thermal vibration. (b) The ligand bound to the active site is stabilized by a P_1 carbonyl oxygen–zinc coordinate bond and other hydrogen bonds with the enzyme. Simultaneously pushing Tyr216 is pushed out of the S_1' pocket. The flexible loop also changes its conformation to make hydrogen bonds with Tyr216 and the ligand. (c) The pushed water molecule is further depolarized by the zinc ion and Glu177 residue, and its nucleophilicity is increased. (d) The water molecule reacts with the carbonyl carbon of the P_1 residue to generate the tetrahedral transition state, which is stabilized by Ala135, Glu177, and Asn191. A proton derived from the water molecule moves from Glu177 carboxyl to -NH- of the scissile bond, (e) and the peptide bind is cleaved. The product is displaced to the 'left' side in the figure, and the side chain of Asn191 is rotated by 94° about the χ_2 axis. As a result, the hydrogen bond between $N\delta_2$–Asn191 and OH–Tyr216 is cleaved, leading to the unstabilization of the fixed flexible loop, and a new hydrogen bond between $N\delta_2$–Asn191 and O–P_3 is then formed. The product is subsequently released from the cleft, partly by the thermal motion of the flexible loop.

O–Ala137, Oδ_1–Glu177, Oδ_1-Asn191 and OH–Try216. The proton the desplaced water molecule is transferred from the carboxyl to the scissile bond to cleave the peptide bond. The product is displaced to the left in the figure, and the side chain of Asn191 is rotated 94° about the χ_2 axis. The hydrogen bond between Nδ_2–Asn191 and OH–Tyr216 is subsequently cleaved, leading to the destabilization of the fixed flexible loop, and a novel hydrogen bond is made between Nδ_2–Asn191 and O–P_3 of the product which is subsequently released from the catalytic cleft by the thermal motion of the flexible loop. As a result, the enzyme is reproducibly set to be ready for the next substrate. When the bound substrate has a basic residue in its P_1 position, a salt bridge may be formed between the side chain of the P_1 residue and O$\delta_{1,2}$–Glu194, and a hydrogen bond between the side chain of the P_1 residue and O–Ala131, leading to the increase of affinity of the ligand to the alkaline protease.

Chapter 4

Acknowledgment

This study was carried out at the Institute for Chemical Research, Kyoto University during the author's Doctor course studies from April 1993 to March 1996. The author is deeply indebted to Associate Professor Yasuo Hata for his continuous guidance, constant encouragement and detailed criticism of this entire manuscript. The author also wishes to acknowledge Professor Sho Takahashi for his general guidance and close collaboration throughout the course of this research, and detailed criticism of the entire manuscript. The author is grateful to Dr. Kazuyuki Morihara for his continuous support for the study and for providing us precious specimens. The author is grateful to Dr. Yukiteru Katsube for his continuous support for this study. The author is also deeply indebted to his co-researchers, Dr. Kensaku Hamada and Mr. Hajime Hiramatsu, for their extremely useful information on the folding topology of *Serratia* protease. The author warmly thanks to Instructor Tomomi Fujii for his devotional cooperation in this study, and would like to express gratitude to all the staff and students of the laboratory. The author is also deeply indebted to Drs. Noriyoshi Sakabe, Atsushi Nakagawa and Nobuhisa Watanabe, National Laboratory of High Energy Physics, for their kind support in the synchrotron experiments. The author would like to express gratitude to all the staff of the Institute for Protein Research, Osaka University, for their kind support in the measurement using the diffractometer. The author is much impressed by Dr. Bunzo Mikami's pertinent advice in data processing. The author thanks Associate Professor Yasushi Akiyama and Instructor Susumu Gotoh for their instruction in using LaTeX. Thanks are also tendered to Miss Martie Jelinek for her proofreading this entire manuscript.

REFERENCES

[1] J. D. Watson and F. H. C. Crick. Molecular structure of nucleic acid. A structure for deoxyribose nucleic acid. *Nature*, 171:737–738, 1953.

[2] J. D. Watson and F. H. C. Crick. Genetic implications of the structure of deoxyribonucleic acid. *Nature*, 171:964–967, 1953.

[3] D. E. McRee. Practical protein crystallography, 1993.

[4] T. L. Blundell and L. N. Johnson. Protein crystallography, 1976.

[5] A. McPherson. Preparation and analysis of protein crystals, 1982.

[6] H. Hope, F. Flow, K. von Böhlen, I. Makowski, C. Kratky, Y. Halfon, H. Danz, P. Webster, K. S. Bartels, H. G. Wittmann, and A. Yonath. Cryocrystallography of ribosomal particles. *Acta Cryst.*, 45:190–199, 1989.

[7] G. A. Petsko. Protein crystallography at sub-zero temperatures: cryoprotective mother liquors for protein crystals. *J. Mol. Biol.*, 96:381–392, 1975.

[8] B. C. Wang. Methods in Enzymology, 1985.

[9] K. Cowtan. *Joint CCP4 and ESF-EACBM Newsletter on Protein crystallography.*, 31:34–38, 1994.

[10] A. Wlodawer, M Miller, M. Jaskolski, B. K. Sathyanarayana, E. Baldwin, I. T. Weber, L. M. Selk, L. Clawson, J. Schneider, and S. B. H. Kent. Conserved folding in retroviral proteases: Crystal structure of a synthetic HIV-1 protease. *Science*, 245:616–626, 1989.

[11] D. W. Cushman, H. S. Cheung, E. F. Sabo, and M. A. Ondetti. Design of potent competitive inhibitors of angiotensin-converting enzyme. carboxyalkanoyl and mercaptoalkanoyl amino acids. *Biochemistry*, 16:5484–5491, 1977.

[12] K. Morihara and H. Tsuzuki. Substrate specificity of elastolytic and nonelastolytic proteinases from *Pseudomonas aeruginosa*. *Arch. Biochem. Biophys.*, 114:158–165, 1966.

[13] K. Morihara and Hiroshige Tsuzuki. *Pseudomonas aeruginosa* peptidohydrolase III. Some characters as a Ca^{2+}-metalloenzyme. *Biochim. Biophys. Acta*, 92:351–360, 1964.

[14] R. B. Honzatko, J. L. Crawford, H. L. Monaco, J. E. Ladner, B. F. P. Edwards, P. F. B. Evans, S. G. Warren, D. C. Wiley, and R. C. Lander W. N. Lipscomb. Crystal and molecular structures of native and CTP-liganded aspartate carbamoyltransferase. *J. Mol. Biol.*, 160:219–263, 1982.

[15] A. Klug and D. Rohdes. "zinc fingers": a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.*, 12:464–469, 1987.

[16] H. Eklund, B. V. Plapp, J. P. Samama, and C. I. Branden. Binding of substrate in a ternary complex of horse liver alcohol dehydrogenase. *J. Biol. Chem.*, 257:14349–14358, 1982.

[17] R. P. J. Williams. The biological chemistry of elements., 1991.

[18] F. Xaver, G. Rüth, F. Grams, I. Yiallouros, H. Nar, U. Küsthardt, R. Zwilling, W. Bode, and W. Stöcher. Crystal structures, spectroscopic features, and catalytic properties of cobalt(II), copper(II), nickel(II), and mercury(II) derivatives of the zinc endopeptidase astacin. *J.Biol. Chem.*, 269:17111–17117, 1994.

[19] D. C. Rees and W. N. Lipscomb. Crystallographic studies on apocarboxypeptidase A and the complex with glycyl-l-tyrosine. *Proc. Natl. Acad. Sci. U. S. A.*, 80:7151–7154, 1983.

[20] G. Shoham, D. C. Rees, and W. N. Lipscomb. Effects of pH on the structure and function of carboxypeptidase A: crystallographic studies. *Proc. Natl. Acad. Sci. U. S. A.*, 81:7767–7771, 1984.

[21] D. W. Christianson and W. N. Lipscomb. Binding of a possible transition state analoque to the active site of carboxypeptidase A. *Proc. Natl. Acad. Sci. U. S. A.*, 82:6840–6844, 1985.

[22] D. W. Christianson and W.N. Lipscomb. X-ray crystallographic investigation of substrate binding to carboxypeptidase A at subzero temperature. *Proc. Natl. Acad. Sci. U. S. A.*, 83:7568–7572, 1986.

[23] G. Shoham, D. W. Christianson, and D. A. Oren. Complex between carboxypeptidase A and a hydrated ketomethylene substrate analoque. *Proc. Natl. Acad. Sci. U. S. A.*, 85:684–688, 1988.

[24] D. W. Christianson, P. R. David, and W. N. Lipscomb. Mechanism of carboxypeptidase A: Jydration of a ketonic substrate analoque. *Proc. Natl. Acad. Sci. U. S. A.*, 84:1512–1515, 1987.

[25] D. W. Christianson and W. N. Lipscomb. Comparison of carboxypeptidase A and thermolysin. *J. Am. Chem. Soc.*, 110:5560–5565, 1988.

[26] D. W. Christianson, L. C. Kuo, and W. N. Lipscomb. Novel structure of the complex between carboxypeptidase A and a ketonic substrate analoque. *J. Am. Chem. Soc.*, 107:8281–8283, 1985.

[27] R. A. Bradshaw, L. H. Ericsson, K. A. Walsh, and H. Neurath. The amino acid sequence of bovine carboxypeptidase A. *Proc. Natl. Acad. Sci. U. S. A.*, 63:1389–1394, 1969.

[28] K. Titani, L. H. Ericsson, K. A. Walsh, and H. Neurath. Amino-acid sequence of bovine carboxypeptidase B. *Proc. Natl. Acad. Sci. U. S. A.*, 72:1666–1670, 1975.

[29] E. Clauser, S. J. Gardell, C. S. Craik, R. J. MacDonald, and W. J. Rutter. Structural characterization of the rat carboxypeptidase A1 and B genes. Comparative analysis of the rat carboxypeptidase gene family. *J. Biol. Chem.*, 263:17837–17845, 1988.

[30] S. J. Gardell, C. S. Craik, E. Clauser, E. J. Goldsmith, C. B. Stewart, M. Graf, and W. J. Rutte. A novel rat carboxypeptidase, CPA2: characterization, molecular cloning, and evolutionary implications on substrate specificity in the carboxypeptidase gene family. *J. Biol. Chem.*, 263:17828–17836, 1988.

[31] D. S. Reynolds, I. L. Stevens, D. S. Gurley, W. S. Lane, K. F. Austen, and W. E. Serafin. Isolation and molecular cloning of mast cell carboxypeptidase A. A novel member of the carboxypeptidase gene family. *J. Biol. Chem.*, 264:20094–20099, 1989.

[32] D. S. Reynolds, D. S. Gurley, R. L. Stevens, D. J. Sugarbaker, K. F. Austen, and W. E. Serafin. Loning of cDNAs that encode human mast cell carboxypeptidase A, and comparison of the protein with mouse mast cell carboxypeptidase A and rat pancreatic carboxypeptidases. *Proc. Natl. Acad. Sci. U. S. A.*, 86:9480–9484, 1989.

[33] D. S. Reynolds, D. S. Gurley, R. L. Stevens, D. J. Sugarbaker, K. F. Austen, and W. E. Serafin. Loning of cDNAs that encode human mast cell carboxypeptidase A, and comparison of the protein with mouse mast cell carboxypeptidase A and rat pancreatic carboxypeptidases. *Proc. Natl. Acad. Sci. U. S. A.*, 86:9480–9484, 1989.

[34] L. D. Fricker, C. J. Evans, F. S. Esch, and E. Herber. Cloning and sequence analysis of cDNA for bovine carboxypeptidase E. *Nature*, 323:461–464, 1986.

[35] W. Gebhard, M. Schube, and M. Eulitz. cDNA cloning and complete primary structure of the small, active subunit of human carboxypeptidase N (kininase 1). *Eur J Biochem*, 178:603–607, 1989.

[36] F. Tan, S. J. Chan, D. F. Steiner, J. W. Schilling, and R. A. Skidgel. Molecular cloning and sequencing of the cDNA for human membrane-bound carboxypeptidase M. Comparison with carboxypeptidases A, B, H, and N. *J. Biol. Chem.*, 264:13165–13170, 1989.

[37] I. Schechter and A. Berger. On the size of the active site in proteases. I. papain. *Biochem. Biophys. Res. Commun.*, 27:157–162, 1967.

[38] W. Stöker, F. Grams, U. Baumann, P. Reinemer, F. Xaver G. Rüth, D. B. Mckay, and W. Bode. The metzincins—topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins collagenases define a superfamily of zinc-peptidases. *Protein Science*, 4:823–840, 1995.

[39] W. Bode, F. X. G. Rüth, and W. Stöckler. Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. *FEBS Lett.*, 331:134–140, 1993.

[40] W. Bode, F. X. Gomis-Rüth, R. Huber, R. Zwilling, and W. Stöcker. Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases. *Nature*, 358:164–167, 1992.

[41] F. Xaver, G. Rüth, L. F. Kress, and W. Bode. First structure of a snake venom metalloproteinase: a prototype for matrix metalloproteinases/collagenases. *Eur. J. Mol. Biol.*, 12:4151–4157, 1993.

[42] W. R. Kester and B. W. Matthews. Crystallographic study of the binding of dipeptide inhibitors to thermolysin: Implications for the mechanism of catalysis. *Biochemistry*, 16:2506–2516, 1977.

[43] M. A. Holmes and B. W. Matthews. Structure of thermolysin refined at 1.6 Å resolution. *J. Mol. Biol.*, 160:623–639, 1982.

[44] A. F. Monzingo and B. W. Matthews. Structure of a mercaptan-thermolysin complex illustrates mode of inhibition of zinc proteases by substrate-analoque mercaptans. *Biochemistry*, 21:3390–3394, 1982.

[45] M. A. Holmes and B. W. Matthews. Binding of hydroxamic acid inhibitors to crystalline thermolysin suggests a pentacoordinate zinc intermediate in catalysis. *Biochemistry*, 20:6912–6920, 1981.

[46] D. E. Tronrud, A. F. Monzingo, and B. W. Matthews. Crystallographic structural analysis of phosphoramidates as inhibitors and transition-state analogs of thermolysin. *Eur. J. Biochem.*, 157:261–268, 1986.

[47] L. H. Weaver, W. R. Kester, and B. W. Matthews. A crystallographic study of the complex of phosphoramidon with thermolysin. a model for the presumed catalytic transition state and for the binding of extended substrates. *J. Mol. Biol.*, 114:119–132, 1977.

[48] D. E. Tronrud, H. M. Holden, and B. W. Matthews. Structures of two thermolysin-inhibitor complexes that differ by a single hydrogen bond. *Science*, 235:571–575, 1987.

[49] F. W. Dahlquist, J. W. Long, and W. L. Bigbee. Role of calcium in the thermal stability of thermolysin. *Biochemistry*, 15:1103–1111, 1976.

[50] A. Molla, Y. Matsumura, T. Yamamoto, R. Okamura, and H. Maeda. Pathogenic capacity of proteases from *Serratia marcescens* and *Pseudomonas aeruginosa* and their suppression by chicken egg white ovomacroglobulin. *Infect. Immun.*, 55:2509–2517, 1987.

[51] A. Molla, T. Yamamoto, T. Akaike, S. Miyoshi, and H. Maeda. Activation of hageman factor and prekallikrein and generation of kinin by various microbial proteinases. *J. Biol. Chem.*, 264:10589–10594, 1989.

[52] A. Molla, S. Tanase, Y. M. Hong, and H. Maeda. Interdomain cleavage of plasma fibronectin by zinc-metalloproteinase from Serratia marcescens. *Biochim. Biophys. Acta*, 955:77–85, 1988.

[53] A. Molla, T. Kagimoto, and H. Maeda. Cleavage of immunoglobulin G (IgG) and IgA around the hinge region by proteases from *Serratia marcescens*. *Infect. Immun.*, 56:916–920, 1988.

[54] A. Molla, K. Matsumoto, I. Oyamada, T. Katsuki, and H. Maed. Degradation of protease inhibitors, immunoglobulins, and other serum

proteins by Serratia protease and its toxicity to fibroblast in culture. *Infect Immun*, 53:522–529, 1986.

[55] K. Matsumoto, T. Yamamoto, R. Kamata, and H. Maeda. Pathogenesis of Serratial infection: Activation of the hageman factor-prekallikrein cascade by Serratial protease. *J. Biochem.*, 96:739–749, 1984.

[56] R. Kamata, T. Yamamoto, K. Matsumoto, and H. Maeda. A Serratial protease causes vascular permeability reaction by activation of the hageman factor-dependent pathway in guinea pigs. *Infect Immun*, 48:747–753, 1985.

[57] A. Molla, T. Yamamoto, T. Akaike, S. Miyoshi, and H. Maeda. Activation of hageman factor and prekallikrein and generation of kinin by various microbial proteinases. *J. Biol. Chem.*, 264:10589–10594, 1989.

[58] H. Maeda, A. Molla, T. Oda, and T. Katsuki. Internalization of Serratial protease into cells as an enzyme-inhibitor complex with alpha 2-macroglobulin and regeneration of protease activity and cytotoxicity. *J. Biol. Chem.*, 262:10946–10945, 1987.

[59] T. Akaike, A. Molla, M. Ando, S. Araki, and H. Maeda. Molecular mechanism of complex infection by bacteria and virus analyzed by a model using Serratial protease and influenza virus in mice. *J. Virol.*, 63:2252–2259, 1989.

[60] K. Morihara. *Pseudomonas aeruginosa* proteinase I. Purification and general properties. *Biochim. Biophys. Acta.*, 73:113–124, 1963.

[61] K. Morihara, H. Tsuzuki, T. Oka, H. Inoue, and M. Ebata. *Pseudomonas aeruginosa* elastase: Isolation, crystallization, and preliminary characterization. *J. Biol. Chem.*, 240:3295–3304, 1969.

[62] M. M. Thayer, K. M. Flaherty, and D. B. Mckay. Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5-Å resoluton. *J. Biol. Chem.*, 266:2864–2871, 1991.

[63] M. M. Smith, L. Shi, and M. Navre. Rapid identification of highly active and selective substrates for stromelysin and matrilysin using bacteriophage display libraries. *J. Biol. Chem.*, 270:6440–6449, 1995.

[64] Y. Shibuya, H. Tanaka, N. Nishino, H. Okabe, T. Kambara, and T. Yamamoto. ctivation of human plasma prekallikrein by Pseudomonas aeruginosa elastase in vitro. *Biochim. Biophys. Acta*, 1097:23–27, 1991.

[65] T. Morimoto, N. Nishino, T. Fujimoto, and T. Yamamoto. Substrate specificity and inhibition of *Pseudomonas aeruginosa* alkaline protease. *Peptide Chemistry*, pages 387–390, 1989.

[66] N. Nishino, W. Shimizu, M. Hirotsuka, T. Fujimoto, and H. Maeda. New synthetic substrates and inhibitors of Serratial 56k protease. *Peptide Chemistry*, 17:2846–2850, 1987.

[67] A. Ducruix and R. Giegeé. Crystallization of nucleic acids and proteins., 1990.

[68] B. W. Matthews. Solvent content of protein crystals. *J. Mol. Biol.*, 33:480–487, 1968.

[69] W. Furey and S. Swaminathan. Am. Crystallogr. Assoc. Mtg. Abst. Ser., 1989.

[70] R. E. Dickerson, F. C. Kendrew, and B. E. Strandberg. The crystal structure of myoglobin: Phase determination to a resolution of 2 å by the method of isomprphous replacement. *Acta Cryst.*, 14:1188–1195, 1961.

[71] N. Sakabe. A focusing Weissenberg camera with multi-layer-line screens for macromolecular crystallography. *J. Appl. Cryst.*, 16:542–547, 1983.

[72] J. Drenth. Principles of protein x-ray crystallography, 1994.

[73] T. Higashi. The processing of diffraction data taken on a screenless weissenberg camera for macromolecular crystallography. *J. Appli. Crystallogr.*, 22:9–18, 1989.

[74] T. A. Jones. Methods in Enzymology, 1985.

[75] Bio-Graphics Inc. *TURBO-FRODO Manual, Version 5.02*, 1994.

[76] K. Matsumoto, H.i Maeda, and R. Okamura. Role of the zinc atom in the activity and conformational stability of Serratial 56K protease. A fluorometric study. *J. Biochem.*, 96:1165–1173, 1984.

[77] R. A. Pauptit, R. Karlsson, D. Picot, J. A. Jenkins, Ann-Sofie Niklaus-Reimer, and J. N. Jansonius. Crystal structure of neutral protease from *Bacillus cereus* refined at 3.0 Å resolution and comparison with the homologous but more thermostable enzyme thermolysin. *J. Mol. Biol.*, 199:525–537, 1988.

[78] K. Morihara, H. Tsuzuki, and T. Oka. On the specificity of *Pseudomonas aeruginosa* alkaline protease with synthetic peptides. *Biochim. Biophis. Acta*, 309:414–429, 1973.

[79] K. Morihara, N. Yoshida, and K. Kuriyama. *Pseudomonas aeruginosa* peptidohydrolase IV. optical rotatory disperion and amino acid composition. *Biochim. Biophis. Acta*, 92:361–366, 1964.

[80] P.E Delepelaire and C. Wandersman. Protease secretion by *Erwinia chrysanthemi*. *J. Biol. Chem.*, 264:9083–9089, 1989.

[81] D. W. Darnall and E. R. Birnbaum. Rare earth metal ions as probes of calcium ion binding sites in proteins. Neodymium (3) acceleration of the activation of trypsinogen. *J. Biol. Chem.*, 245:6484–6486, 1970.

[82] A. T. Brünger. X-plor, 1992.

[83] A. T. Brünger, J. Kuriyan, and M. Karplus. Crystallographic R factor refinement by molecular dynamics. *Science*, 235:458–460, 1987.

[84] A. T. Brünger and A. Krukowski. Slow-cooling protocols for crystallographic refinement by simmulated annealing. *Acta Cryst.*, A46:585–593, 1990.

[85] R. A. Engh and R. Huber. Accurate bond and angle parameteres for X-ray protein-structure refinement. *Acta Cryst.*, A47:392–400, 1991.

[86] A. T. Brünger. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 355:472–475, 1992.

[87] P. V. Luzzati. Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Cryst.*, 5:802–810, 1952.

[88] A. J. C. Wilson. The probability distribution of X-ray intensities. *Acta Cryst.*, 2:318–321, 1949.

[89] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26:283–291, 1993.

[90] W. Kabsch and C. Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22:2577–2637, 1983.

[91] N. S. Templeton, P. D. Brown, A. T. Levy, I. M. Margulies, L. A. Liotta, and W. G. S. Stevenson. Cloning and characterization of human tumor cell interstitial collagenase. *Cancer Res.*, 50:5431–5437, 1990.

[92] E. Dumermuth, J. A. Eldering, J. Gruenberg, W. Jiang, and E. E. Sterchi. Cloning of the PABA peptide hydrolase alpha subunit (PPH-α) from human small intestine and its expression. *FEBS Lett.*, 335:367–375, 1993.

[93] H. M. Holden, D. E. Tronrud, A. F. Monzingo L. H. Weaver, and B. W. Matthews. Slow- and fast-binding inhibitors of thermolysin display different modes of binding: crystallographic analysis of extended phosphonamidate transition-state analogues. *Biochemistry*, 26:8542–8553, 1987.

[94] H. Takeya, M. Arakawa, T. Miyata, and S. Iwanaga. Primary structure of H-2-proteinase, a non-hemorrhagic metalloproteinase, isolated from the venom of the habu snake, trimeresurus flavoviridis. *J. Biochem.*, 106:151–157, 1989.

[95] G. J. Barton. ALSCRIPT a tool to format multiple sequence alignments. *Protein Eng.*, 6, 1993.

[96] M. Yoder, N. T. Keen, and F. Jurnak. New domain motif: The structure of pectate lyase c, a secreted plant virulent factor. *Science*, 260:1503–1507, 1993.

[97] M. Yoder, S. E. Lietzke, and F. Jurnak. Unusual structural features in the parallel β-helix in pectate lyase. *Structure*, 1:241–251, 1993.

[98] S. Steinbacher, R. Seckler, S. Miller, B. Steipe, R. Huber, and P. Reinemer. Crystal structure of P22 tailspike protein: Interdigitated subunits in a thermostable trimer. *Science*, 265:383–386, 1994.

[99] B. W. Matthews, J. N. Jansonius, P. M. Colman, B.P. Schoenborn, and D. Dupourque. Three-dimensional structure of thermolysin. *Nature New Biol.*, 238:37–41, 1972.

[100] P. M. Colman, J. N. Jansonius, and B. W. Matthews. The structure of thermolysin: an electron density map at 2-3 Å resolution. *J. Mol. Biol.*, 70:701–724, 1972.

[101] B. W. Matthews. Structural basis of the action of thermolysin and related zinc peptidases. *Acct. Chem. Res.*, 21:333–340, 1988.

[102] B. L. Vallee and D. S. Auld. Active-site zinc ligands and activated H_2O of zinc enzymes. *Proc. Natl. Acad. Sci. U.S.A.*, 87:220–224, 1990.

[103] K. Nakahama, K. Yoshimura, R. Marumoto, M. Kikuchi, I. S. Lee, T. Hase, and H. Matsubara. Cloning and sequencing of *Serratia* protease gene. *Nucleic Acids Res.*, 14:5843–5855, 1986.

[104] U. Baumann. Crystal structure of the 50 kDa metallo protease from *Serratia marcescens*. *J. Mol. Biol.*, 242:244–251, 1994.

[105] F. X. Gomis-Rüth, W. Stöker, R. Huber, R. Zwilling, and W. Bode. Refined 1.8 Å X-ray crystal structure of astacin, a zinc-endopeptidase from the crayfish *Astacus astacus* l. *J. Mol. Biol.*, 229:945–968, 1993.

[106] B. Lovejoy, A. Cleasby, A. M. Hassell, K. Longley, M. A. Luther, D. Weigl, G. McGeehan, A. B. McElroy, D. Drewry, M. H. Lmbert, and S. R. Jordan. Structure of the catalytic domain of fibrobrast collagenase complexed with an inhibitor. *Science*, 263:375–377, 1994.

[107] J. C. Spurlino, A. M. Smallwood, D. D. Carlton, T. M. Banks, K. J. Vavra, J. S. Johnson, E. R. Cook, J. Falvo, R. C. Wahl, and T. A. Pulvino. 1.56 Å structure of mature truncated human fibroblast collagenase. *Proteins*, 19:98–109, 1994.

[108] F. X. G. Rüth, J. Kellermann L. F. Kress, I. Mayr, X. Lee, R. Huber, and W. Bode. Refined 2.0 Å X-ray crystal structure of the snake venom zinc-endopeptidase adamalysin II. primary and tertiary structure determination, refinement, molecular structure and comparison with astacin, collagenase and thermolysin. *J. Mol. Biol.*, 239:513–544, 1994.

[109] P. Delepelaire and C. Wandersman. Protese secretion by *Erwinia chrysanthemi*. *J. Biol. Chem.*, 264:9083–9089, 1989.

[110] H. Takagi H. Ikemura and M. Inouye. Requirement of pro-sequence for the production of active subtilisin e in Escherichia coli. *J. Biol. Chem.*, 262:7859–7864, 1987.

[111] W. Sücker, M. Ng, and D. S. Auld. Fluorescent oligopeptide substrates for kinetic characterization of the specificity of *Astacus* protease. *Biochemistry*, 29:10418–10425, 1990.

[112] T. Ogita, A. Sato, R. Enokita, K. Suzuki, M. Ishii, T. Negishi, T. Okazaki, K. Tamaki, and K. Tanzawa. Matlystatins, new inhibitors of type IV collagenases from *Actinomadura atramentaria*. *J. Antibiotics*, 45:1723–1732, 1992.

[113] S. N. Arnett, G. Fields, H. Birkedal-Hansen, and H. E. V. Wart. Sequence specificities of human fibroblast and neutrophil collagenases. *J. Biol. Chem.*, 266:6747–6755, 1991.

[114] K. Imai, Y. Yokohama, I. Nakanishi, E. Ohuchi, Y. Fujii, N. Nakai, and Y. Okada. Matrix metalloproteinase 7 (matrilysin) from human rectal carcinoma cells. *J. Biol. Chem.*, 270:6691–6697, 1995.

[115] M. F. Browner, W. W. Smith, and A. L. Castelhano. Matrilysin–inhibitor complexes: common themes among metalloproteases. *Biochemistry*, 34:6602–6610, 1995.

[116] J. Cao, H. Sato, T. Takino, and M. Seiki. The C-terminal region of membrane yype matrix metalloproteinase is a functional transmembrane domain required for pro-gelatinase A activation. *J. Biol. Chem.*, 270:801–805, 1995.

[117] A. Y. Strongin, I. Collier, G. Bannikov, B. L. Marmer, G. A. Grant, and G. I. Goldberg. Mechanism of cell surface activation of 72-kDa type IV collagenase. *J. Biol. Chem.*, 270:5331–5338, 1995.

[118] T. N. Young, Salvatore V. Pizzo, and M. Sharon Stack. A plasma membrane-associated component of ovarian adenocarcinoma cells enhances the catalytic efficiency of matrix metalloproteinase–2. *J. Biol. Chem.*, 270:999–1002, 1995.

[119] Petra marchand, J. Tang, G. D. Johnson, and J. S. Bond. Coohterminal proteolytic processing of secreted and membrane forms of the α subunit of the metalloprotese subunit of the metalloprotease meprin A. *J. Biol. Chem.*, 270:5449–5456, 1995.

[120] W. Stöcher, F. X. Gomis-Rüth, W. Bode, and R. Zwilling. Implications of the three-dimensional structure of astacin for the structure and function of the astacin family of zinc-endopeptidases. *Eur. J. Biochem.*, 214:215–231, 1993.

[121] F. Duong, A. Lnzdunski, B. Cami, and M. Murgier. Sequence of a cluster of genes controlling synthesis and secretion of alkaline protease in *Pseudomonas aeruginosa*: relatonships to other secretory pathways. *Gene*, 121:47–54, 1992.

[122] J. Guzzo, M. Murgier, A. Filloux, and A. Lazdunski. Cloning of the *Pseudomonas aeruginosa* alkaline protease gene and secretion of the protease into the medium by *Escherichia coli*. *J. Bacteriol.*, 172:942–948, 1990.

[123] V. Koronakis, M. Cross, B. Senior, E. Koronakis, and C. Hughes. The secreted hemolysins of *Proteus mirabilis, Proteus vulgaris*, and *Morganella morganii* are genetically related to each other and to the alpha-hemolysin of *Escherichia coli*. *J. Bacteriol.*, 169:1509–1515, 1987.

[124] N. Mackman, J.-M. Nicaud, L. Cray, and I. B. Holland. Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*. *Mol. Gen. Genet.*, 201:529–536, 1985.

[125] P. Delepelaire and C. Wandersman. Protein secretion in gram-negative bacteria. the extracellular metalloprotease B from *Erwinia chrysanthemi* contains a C-terminal secretion signal analogous to that of *Escherichia coli* alpha-hemolysin. *J. Biol. Chem.*, 265:17118–17125, 1990.

[126] U. Baumann, M. Bauer, S. Létoffè, P. Delepelaire, and C. Wandersman. Crystal structure of a complex between *Serratia marcescens* metalloprotease and an inhibitor from *Erwinia chrisanthemi*. *J. Mol. Biol.*, 248:653–661, 1995.

[127] U. Baumann, S. Wu, K. M. Flaherty, and D. B. Mckay. Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif. *EMBO J.*, 12:3357–3364, 1993.

[128] A. Ludwig, T. Jarchau, R. Benz, and W. Goebel. The repeat domain of *Escherichia coli* haemolysin (hlya) is respensible for its Ca^{2+}-dependent binding to erythrocytes. *Mol. Gen. Genet.*, 214:553–561, 1988.

[129] H. Miyatake, Y. Hata, T. Fujii, T. Akutagawa, K. Hamada, and Y. Katsube. Crystal structure of aklaline protease from *Pseudomonas aeruginosa* IFO3455. *Bulletin of the Institute for Chemical Research, Kyoto Univ.*, 72:373–386, 1994.

[130] H. Miyatake, Y. Hata, T. Fujii, K. Hamada, K. Morihara, and Y. Katsube. Crystal structure of unliganded alkaline protease from *Pseudomonas aeruginosa* IFO3080 and its conformational changes on ligand binding. *J. Biochem.*, 118:474–479, 1995.

[131] J. Hajdu, P. A. Machin, J. W. Campbell, T. J. Greenhough, I. J. Clifton, S. Zurek, S. Gover, L. N. Johnson, and M. Elder. Millisecond X-ray diffraction and the first electron density map from laue photographs of a protein crystal. *Nature*, 329:178–181, 1987.

[132] I. Schlichting, S. C. Almo, G. Rapp, K. Wilson, K. P., and R. S. Goody. Time-resolved X-ray crystallographic study of the conformational change in Ha-Ras p21 protein on GTP hydrolysis. *Nature*, 345:309–315, 1990.

[133] P. Carloni S. Mangani and P. Oriol. Crystal structure of the complex between carboxypeptidase a and the biproduct analog inhibitor l-benzylsuccinate at 2.0 Å resolution. *J. Mol. Biol.*, 223:573–578, 1992.

[134] H. Kim and W. N. Lipscomb. Crystal structure of the complex of carboxypeptidase a with a strongly bound phosphonate in a new crystalline form: comparison with structures of other complexes. *Biochemistry*, 29:5546–5555, 1990.

[135] D. W. Christianson and W. N. Lipscom. X-ray crystallographic investigation of substrate binding to carboxypeptidase a at subzero temperature. *Proc. Natl. Acad. Sci. U. S. A.*, 83:7568–7572, 1986.

[136] W. Bode, P. Reinemer, R.t Huber, T. Kleine, S. Schnierer, and H. Tscheshe. The X-ray crystal structure of the catalytic domain of human neutrophil collagenase inhibited by a substrate analogue reveals the essentials for catalysis and specificity. *Eur. J. Mol. Biol.*, 13:1263–1269, 1994.

[137] M. G. Rossmann. Processing oscillation data for very large unit cells with an atomatic convolution technique and profile fitting. *J. Appl. Cryst.*, 12:225–238, 1979.

[138] M. G. Rossmann and D. M. Blow. The detection of sub-units within the crystallographic asymmetric unit. *Acta Cryst.*, 15:24–31, 1962.

[139] A. T. Brünger. Extension of molecular replacement: A new search strategy based on patterson correlation refinement. *Acta Cryst.*, A46:46–57, 1990.

[140] *Quanta version 3.3 Parameter Handbook.* NewYork, U. S. A., Molecular Simulations, Inc. edition, 1992.

[141] B. William, N. Lipsom, Jean A. Hartsuck, F. A. Quiocho, and Jr. G. N. Reeke. The structure of carboxypeptidase A, IX. The X-ray diffraction results in the light of the chemical sequence. *Proc. Natl. Acad. Sci. U. S. A.*, 64:28–35, 1969.

[142] S. J. Gardell, C. S. Craik, D. Hilvert, and M. S. Urdea W. J. Rutter. Site-directed mutagenesis shows that tyrosine-248 of carboxypeptidase A does not play a crucial role in catalysis. *Nature*, 317:551–555, 1985.

[143] J. T. Johansen and B. L. Vallee. Environment and conformation dependent sensitivity of the arsanilazotyrosine-248 carboxypeptidas a chromophore. *Biochemistry*, 14:649–660, 1975.

[144] P. Chakrabarti. Geometry of interaction of metal ions with sulfur-containing ligands in protein structures. *Biochemistry*, 28:6081–6085, 1995.

[145] F. Grans, P. Reinemer, J. C. Powers, T. Kleine, M. Pieper, G. Tschesche, R. Huber, and W. Bode. X-ray structures of human neutrophil collagenase complexed with peptide hydroxamate and peptide thiol inhibitors. Implications for substrate binding and rational drug design. *Eur. J. Biochem.*, 228:830–841, 1995.

[146] F. X. G. Rüth, F. Grams, I. Yiallouros, H. Nar dnd U. Kusthardt, R. Zwilling, W. Bode, and W. Stocker. Crystal structures, spectroscopic features, and catalytic properties of cobalt(II), copper(II), nickel(II), and mercury(II) derivatives of the zinc endopeptidase astacin. A correlation of structure and proteolytic activity. *J. Biol. Chem.*, 269:17111–17117, 1994.

[147] Y. Shibuya, T. Yamamoto, T. Morimoto, N. Nishino, T. Kambara, and H. Okabe. *Pseudomonas aeruginosa* alkaline protease might share a biological function with plasmin. *Biochim. Biophis. Acta*, 1077:316–324, 1991.